VOL. 2, NO. 1, FEBRUARY, 1960

human factors

THE JOURNAL OF THE HUMAN FACTORS SOCIETY

Contents

HUMAN FACTORS IN HIGHWAY DESIGN, OPERATION AND SAFETY PROBLEMS 1
Theodore W. Forbes

PROSPECTUS FOR HUMAN FACTORS IN ELECTRONICS 9
Elvin E. Herman

A COMPARISON OF THREE FLIGHT ATTITUDE DISPLAYS 14
Raymond E. Bernberg

SPEECH INTERFERENCE LEVEL AND AIRCRAFT ACOUSTICAL ENVIRONMENT 18
John J. Dreher and William E. Evans

OPTIMUM SYMBOLS FOR RADAR DISPLAYS 28
Hugh M. Bowen, John L. Andreassi, Shaffer Truax and Jesse Orlansky

OPERATOR TRENDS IN MARKING A MOVING SCALE INDICATOR 34
Harold R. Willis

BOOK REVIEW 44

THE AUTHORS 46

EDITORIAL 49

PERGAMON PRESS

OFFICERS AND EXECUTIVE COUNCIL

T58
H3 H8
v.2, 1960-61

The Human Factors Society is a non-profit California Corporation. The officers and members of the Executive Council are:

PRESIDENT: Arnold M. Small, 1960
PRESIDENT-ELECT: Stanley N. Roscoe, 1960
VICE-PRESIDENT: Smith W. Ames, 1960
PAST-PRESIDENT: Renato Contini, 1960
SECRETARY: Paul G. Cheatham, 1960
TREASURER: Donald W. Conover, 1960
EDITOR-IN-CHIEF: Stanley Lippert, 1959-1961

EXECUTIVE COUNCIL:

Charles I. Barron, 1960-1962
Paul M. Fitts, 1959-1960
Jack A. Kraft, 1959-1960
Max W. Lund, 1959-1961
John Lyman, 1960-1962
Ross A. McFarland, 1959-1961

EDITORIAL BOARD:

Walter P. Grether
Blaine H. Levedahl
J. C. R. Licklider
Stanley Lippert
John Lyman
John P. Meehan
Allison K. Simons
Ezra V. Saul
Franklin V. Taylor
Neil D. Warren
Joseph W. Wulfeck

HISTORIAN: Donald T. Hanifan

COPYRIGHT ©, 1960
by the HUMAN FACTORS SOCIETY

by PUBLISHED BY
PERGAMON PRESS, 122 E. 55th St., New York 22, N. Y.
PE 4/5 Fitzroy Square, London W. 1. England.
Printed in U.S.A.

Address general correspondence to Human Factors Society, P. O. Box 263, Arlington 10, Virginia, Attention: **Paul G. Cheatham**, Secretary, for Society business including payment of dues.

Address matters pertaining to the Journal or Bulletin to **Human Factors**, Biotechnology Laboratory, Department of Engineering, University of California, Los Angeles 24, California, Attention:

John Lyman, Managing Editor, for Journal or Bulletin Subscriptions or other publication business.

Stanley Lippert, Editor-in-Chief, for submission of Journal articles and letters to the editor.

Hilde Groth, Bulletin Editor, for Bulletin news.

SUBSCRIPTION PRICE MEMBERSHIP in Society includes subscription to both publications. Journal. NON-MEMBERS SUBSCRIPTIONS TO BE SENT DIRECT TO THE PUBLISHER, PERGAMON PRESS. Commencing with Volume 2 the subscription price will be: $10.00 (£3.10s.) per annum, for individuals who certify that the journal is for their personal use. $20.00 (£7) per annum, for libraries, government establishments and research institutions.

Bulletin, $5.00 per year. Orders to be sent to John Lyman.

Human Factors in Highway Design, Operation and Safety Problems[1]

THEODORE W. FORBES, Highway Traffic Safety Center, Michigan State University, East Lansing, Michigan.

On several urban freeways, peak hour traffic leads drivers to operate vehicles at time spacing less than their fastest emergency reaction time. Therefore, factors affecting speed and accuracy of driver judgment and response have become critically important. Experimental studies have shown how driver perception, judgment, and response time (measured by driver lag in acceleration and deceleration) can be affected by highway design features. Experimental studies of factors affecting driver responses when combined with mathematical analyses based on such driver-vehicle-highway responses may explain differences in maximum traffic flow at which self limiting stoppage occurs without adequate physical causes in heavy traffic.

Other human engineering data already available or obtainable through specialized studies suggest ways of obtaining more efficient driver-vehicle-highway-and-environment interrelationships and, thereby, more efficient and safer operation of highways. Examples of such information and studies are given.

Importance in Other Areas

Human reactions have been widely recognized as being of importance in safety education and enforcement fields of highway and traffic safety. In these areas attitudes, motivation of drivers, learning of operating procedures and the like are of primary importance. These may be viewed as the direct action phases of the approach to highway safety problems. Millions of dollars and untold hours of manpower have been spent in these areas, but we still do not know which techniques are most effective and whether better ones may not be developed. Much research is underway in this area, and much more is needed.

Importance in Highway Design and Operation Problems

Less recognized is the fact that human factors have now become of major importance in design and operation problems of highways both for efficient traffic flow and for safety. With the increased number of vehicles and more free flowing highways of modern design, the relation of operator, vehicle and environment (including highway, weather and other phases) have become critically important. In earlier years, it was sufficient to build a highway that would stand up under loads and then leave it for anyone to use as he saw fit. Now, however, with the increased vehicle volume on the highways and the more reliable and higher performance of vehicles, it has become clear that both traffic engineers and those designing the geometrics of the highway must plan for highway operation on a scientific basis.

Perception, Judgment, and Response Time Now Vital

On our urban freeways (divided highways with grade-separated interchanges and controlled access from the sides) the pressure of traffic at peak hours leads to driver performance at time spacings where perception-judgment-and-response time requirements are of critical magnitude. Where two or three drivers have the highway to themselves, this may not be true. However, where multiple lane freeways carry 1800 to 2400 cars per hour in a single lane, such time factors are very critical. At these rates of flow, time headways *average* from one and one-half to two seconds. However, a study of headways on the Pasadena Freeway showed that *individual headways within platoons* under peak conditions consistently approached one-half second (Fig. 1) at almost 50 miles per

[1] Presented at the Annual Meeting of the Human Factors Society, Los Angeles, 17–18 September 1959.

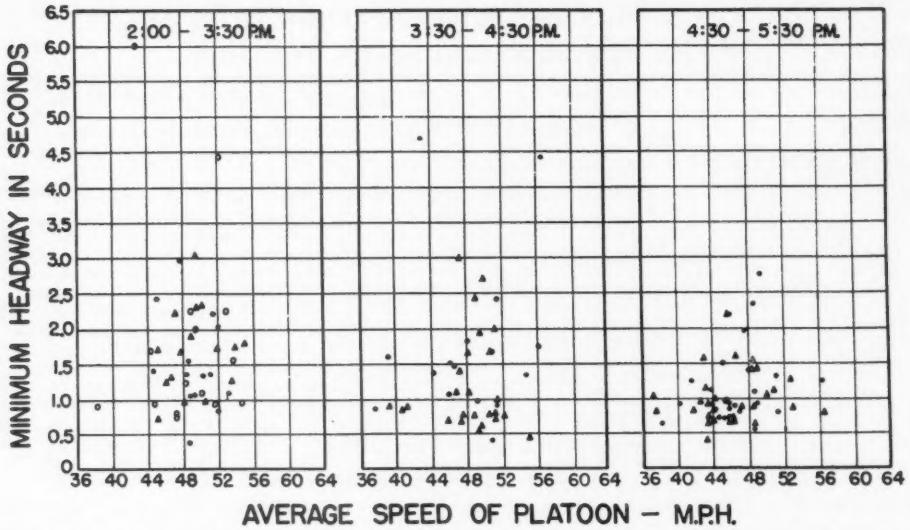

Fig. 1. Individual headways approached 0.5 seconds in heavy traffic. Minimum headway and average speed of platoon — Lane 2 for light, medium, and heavy traffic periods. (From Reference 3.)

hour[3]. This represented approximately 35 feet center to center between cars of which approximately one-half represented the length of the vehicle. Therefore, the clear space from the front bumper to the rear bumper of the car ahead represented approximately one-quarter second.

Data from a recent study at our Center corroborated the earlier study. Again individual time headways within platoons at one location on a Detroit freeway approached 0.5 seconds in peak traffic. (Fig. 2) The *average headway* for this 15 minute time period at this location was 1.5 seconds. At a location on another Detroit expressway, average headways of 1.5 to 2.0 seconds were reached at between 40 and 50 miles per hour[11] (see especially Section V and VI of Reference 11), but this average included some of the larger gaps between platoons. Within platoons, a significant proportion of individual drivers operated under the shorter and much more critical time response limitations.

This means that drivers must have based their judgments on the maneuvers of a vehicle several cars ahead. Successful driving under these conditions required vehicles to follow each other at a time spacing which is approximately one-half the fastest driver response time for alert, emergency conditions, and yet drivers are doing this (for the most part successfully) every day in Los Angeles and other large cities. Hence the importance of human response time is clear.

Limitation of Traffic Flow by Human and Vehicle Response Time

For many years it has been known that at certain maximum volumes of flow, highway traffic develops waves of temporary stoppage without any apparent physical reason. This self-limiting phenomenon occurs on freeways as well as on ordinary highways and has been measured and studied in various vehicle tunnels in New York City[2]. There it occurs even while traffic is fed steadily into and flows freely out of the ends of the tunnels. Traffic measurements showed that this self-limiting stoppage occurs at different maximum flow in different tunnels and other studies suggest that it occurs at different maxima on different freeways.

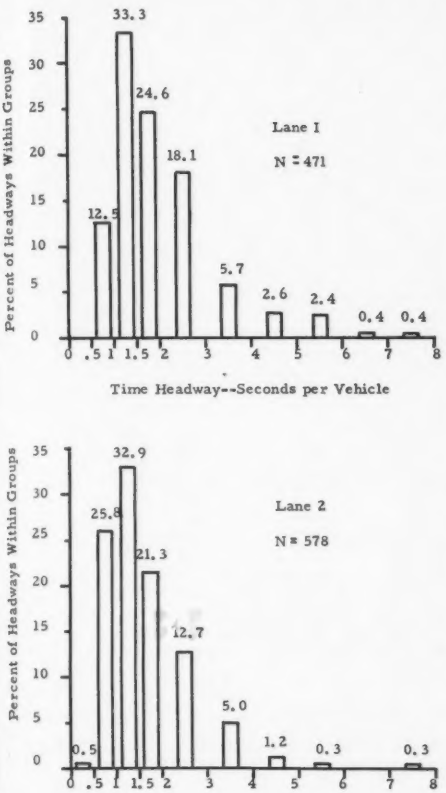

Fig. 2. Distribution of individual time headways. Ford Expressway, Detroit, 7:15–7:30 A.M. 3–27–58.
Lane 1 — Inside lane next to medial divider, Lane 2 — Middle lane, of six-lane freeway. Data analyzed specially for this report.

A possible theoretical explanation for "self bottlenecking" at different volumes can be drawn from human engineering data available from various studies. For instance, when various studies were put together and interrelated, it was clearly shown that the time required for human perception-judgment-and-response increased consistently with the number of stimuli and the number of responses required[6], and the time required increased consistently with the complexity of the operator's task in relation to the vehicle and the roadway environment (Figs. 3 and 4).

Factors Which May Be Involved

The vehicle-highway-environment factors which may increase the complexity of the operator's task are sometimes obvious ones like more controls to manipulate, sometimes less obvious ones (such as an unnoticed reduction of visibility), and sometimes even less obvious ones than these.

Interest in different maximum volume before self-limitation in different tunnels led to an experimental study[8] of operator reactions using an experimental three-car platoon. Photographic records were made of driver responses in cars 2 and 3 to sudden deceleration and acceleration of the first vehicle at certain predetermined locations. The use of identical, high performance, light vehicles approximately equated vehicle lags. Resulting driver response lags in reaccelerating were measured and plotted. The results showed increased driver lags with decreased visibility and lighting conditions, on a down grade, on a right curve and with a situation involving psychological constriction. (Figs. 5 and 6)

These increases in driver lag were of sufficient magnitude to explain different maximum flow levels in different tunnels.

Relation to Mathematical Models and Simulation of Flow

For several years, interest has been increasing in the formulation of mathematical models looking toward simulation of traffic flow. Such mathematical simulation with the rapidly improving electronic computers can be of great importance for testing out traffic problems.

However, a very complex phenomenon such as highway traffic is difficult to simulate accurately unless all important factors are included. Of utmost importance are the human operator-vehicle-highway-environment relationships. Possibly for some purposes it may be sufficient to assume that the operator's reaction averages out. However, in the case of freeway flow and tunnel flow this has not been satisfactory.

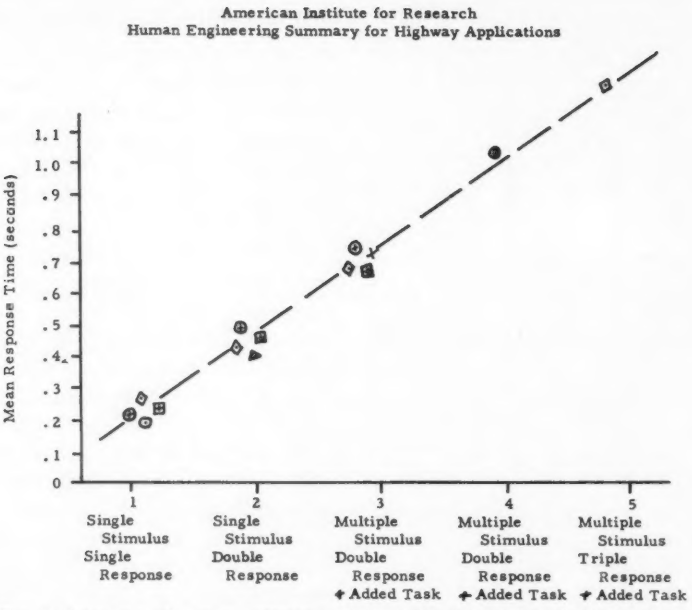

Fig. 3. Increase of mean response time with increasing tasks and responses. Data from six studies of foot and leg responses on brake and accelerator; one of tracking with control stick. (See Reference 6, p. 17.)

A mathematical model based on fluid flow theory has been compared to actual traffic flow data both on freeways and in tunnels. This mathematical model is satisfactory to a point but does not fit the data as accurately as one would desire[2, 9, 11]. Furthermore, as already mentioned, maximum volume, density and speed values proved to be different for several essentially similar highways and vehicular tunnels. This particular approach assumed that such different flow maxima are characteristic of different highway facilities but did not furnish a logical basis for such different maxima, nor for human factor variables.

Another analysis on the basis of a servomechanical analogy[1, 10] included a factor representing driver sensitivity of judgment and response. A very similar mathematical function relating traffic flow, density and speed was developed without the somewhat illogical assumption of fluid flow. Experimental measurement of a driver sensitivity factor by means of an experimental two-car follow-the-leader setup indicated a possible velocity dependence or distance-headway dependence of this sensitivity factor in level, open driving.

Combination of the Two Approaches

The driver lags from the earlier three-car platoon study[8] provided data on driver sensitivity of a type related to that used by Herman[1]. The greater range of experimental variables gave a somewhat greater range of lag times than those measured in his two-car experiment in open, level driving. This greater number of variables was easily handled in the driving experiment but may pose greater difficulties for the mathematician. These two approaches viewed together can account logically for stagnation at different levels of flow, on different highway facilities. Such an explanation also offers possibilities of increasing maximum traffic flow on highways through study of the pertinent psychological factors affecting operators so as to allow their introduction into mathematical models and simulation.

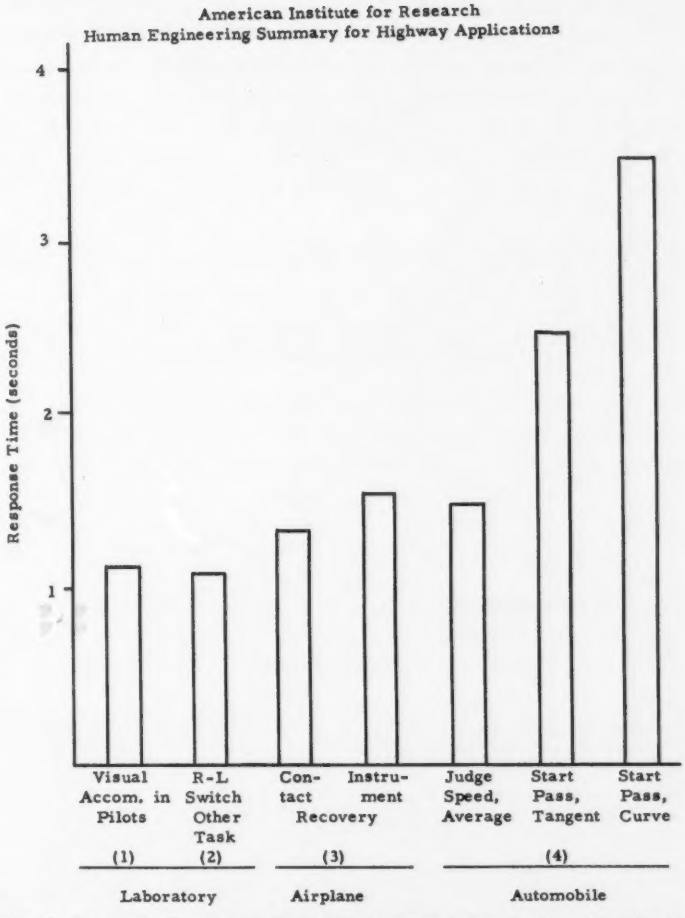

Fig. 4. Increase of response time with increased complexity and difficulty of judgment task. (See Reference 6, p. 23.)

Other Applications of Human Factors Data

For other applications, also, data are available[4] showing how relationships between human sensitivity and reaction can be analyzed in a useful fashion. For instance, data on human vision were used to illustrate how design of tunnel entrances and lighting can take into account human reactions to eliminate the effect of driving into a dark hole. This effect showed up as an important influence on driver reactions in the three-car platoon experiment.

Familiarity With the Situation — Preparatory Information

Reduction of operator response time and reduction of errors can be accomplished (1) by simplifying the requirements of the task and (2) providing previous understanding and prepared responses on the part of operators. The first often can be provided by proper human engineering design, for example, the illustration of the tunnel entrance, or in interchange design, by reducing the number of choices at any given ramp or junction point. The second

Fig. 5. Driver lag increase with poorer lighting-and-visibility. Response lag in re-accelerating. (See Reference 8, Fig. 3.)

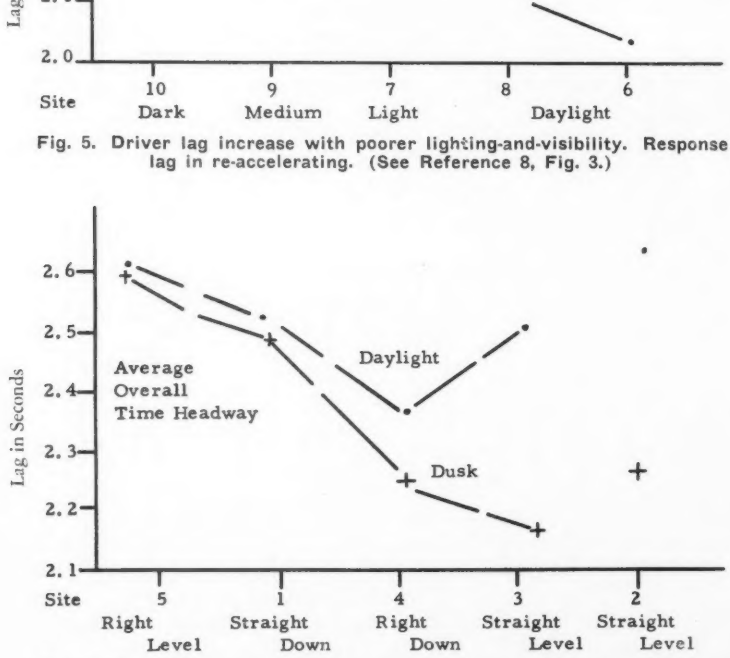

Fig. 6. Driver lag increase with other factors. Response lag in re-accelerating. (See Reference 8, Fig. 2.)

can be provided by advance information from maps and booklets at service stations and by improved sign design on the highway, combined with driver education.

A number of studies have provided sign design data which are now available to traffic engineers and others designing highway signs. (For example, see Reference 7.) These data allow the design of signs which can be read at required distances which should be calculated in terms of seconds of warning time using the 85 percentile speed on the highway.

One paper reported observations which may be corroborated by almost anyone showing motorists who suddenly changed their minds and veered across the highway at a junction, thus creating considerable hazard[4]. Such hazards can be reduced or eliminated by better sign design and better pre-information.

A recent study[12] investigated the kind of information needed by motorists who use the signs on a freeway. Very significantly this study showed that the motorists who needed the signs most were the relatively small group of strange drivers unfamiliar with the location. In some cases, in spite of supposedly good freeway signing, they did not realize that they had taken a wrong turn and had arrived at an undesired location.

Self-Explanatory Symbols Needed

Other research can contribute better methods of giving the driver the most preparation and the most useful and needed information quickly and effectively. One example is the use of more self-explanatory symbols for highway signs and signals.

A recent study of symbols for individual lane control signals[5] on multi-lane highways or bridges was carried out to eliminate the undesirable use of a stop signal which does not mean "stop." In a number of different cities, individual signals have been used to control separately each lane of a multi-lane highway or tunnel. Such independent control may be needed in case of an accident or to reverse flow in certain lanes where the inbound and outbound load is unbalanced during morning and evening peak traffic. Standard red-bullseye "stop" signals are often used for lane control and also at toll booths. At ordinary street intersections the signal means "stop here" but in other cases, when a toll booth is closed, the signal means "do not use this entrance." Or, in case of an accident or breakdown where it is desired to bring in a tow-truck, it means "do not use this lane." It has been found on some of these facilities that drivers obey the signal by stopping, thus clogging the lane until a police officer is sent to extricate them.

An engineering psychology study was carried on in the laboratory of a series of symbols. A pictorial presentation of the symbols was given in three different ways to several groups of subjects. A red-X symbol showed the largest overall proportion of "do not use" or "do not travel in this lane" interpretations and the least proportion of "stop right here." The combination of this red-X and a green-arrow-pointing-up (to distinguish it from a green-arrow-pointing-down which is often used on signs indicating "this lane for a given destination") are now in use on the Mackinac Bridge in Michigan. (Fig. 7.)

To double check the laboratory findings, the actual responsiveness of drivers to the red-X signal was checked by a simple experimental

Fig. 7. View of signals on Mackinac Bridge showing design of red-X, green arrow and speed signals. (From Reference 5.)

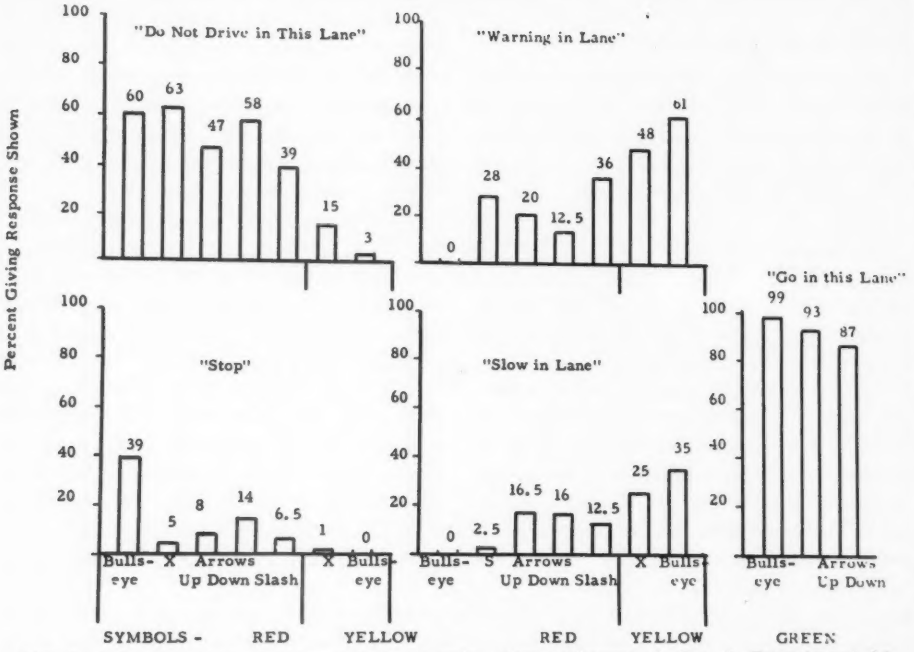

Fig. 8. Interpretations of lane control symbols. (See Reference 5, Fig. 1, Experiment 1.)

setup in actual traffic on the bridge. This study used a light wooden barrier to eliminate any possible hazard if it was struck. Drivers responded to the red-X by moving over significantly farther ahead of the barrier than they did when the green arrow was showing (Fig. 8).

CONCLUSION

Thus proper attention to human engineering factors in design and operation of highways can contribute greatly to further improvement of traffic efficiency and safety. Much applicable human engineering information is available and techniques for obtaining more are available.

REFERENCES

1. CHANDLER, R. E., HERMAN, R., and MONTROLL, E. W., Traffic dynamics: Studies in car following. *Operations Research,* **6**, 165–184, 1958.

2. EDIE, L. C., and FOOTE, R. S., Traffic flow in tunnels. *Highway Research Board,* January, 1958.

3. FORBES, T. W., Speed, headway and volume relationships on a six-lane freeway. *Proc. Inst. Traf. Engrs.,* 103–126, 1951.

4. FORBES, T. W., Driver characteristics and highway operation. *Traffic Engineering,* **24**, 49–51, 1953. (Also *Proc. Inst. Traf. Eng.,* 68–73, 1953.)

5. FORBES, T. W., GERVAIS, E., and ALLEN, TERRENCE, Effectiveness of symbols for lane control signals. *Highway Research Board,* 1959. (In Press.)

6. FORBES, T. W., and KATZ, M. S., Summary of human engineering research data and principles related to highway design and traffic engineering problems. American Institute for Research, Pittsburgh, Pa., pp. 110, April, 1957.

7. FORBES, T. W., MOSKOWITZ, K., and MORGAN, G., A comparison of lower case and capital letters for highway signs. *Proc. Highway Research Board,* **30**, 355–373, 1950. (Also in California Highways and Public Works Magazine, **61**, 6–9, 1951.)

8. FORBES, T. W., ZAGORSKI, H. J., HOLSHOUSER, E. L., and DETERLINE, W. A., Measurement of driver reactions to tunnel conditions and effects on traffic flow. *Highway Research Board,* Bull. No. 37, 345–357, 1958.

9. GREENBERG, H., Tunnel traffic capacity study. *Operations Research,* **7**, 79–85, 1959.

10. HERMAN, R., MONTROLL, E. W., POTTS, R. B., and ROTHERY, R. W., Traffic dynamics: Analysis of stability in car following. *Operations Research,* **7**, 86–106, 1959.

11. MAY, ADOLF D. JR., Traffic characteristics and phenomena on high density controlled access facilities, presented at the Los Angeles Conference of the American Society of Civil Engineers, February 12, 1959.

12. SCHOPPERT, D. W., MOSKOWITZ, K., HULBERT, S. F., and BURG, A., Some principles of freeway directional signing based on motorists' experiences, 38th Annual Meeting, *Highway Research Board,* Washington, D.C., January, 1959. (In Press.)

Prospectus for Human Factors in Electronics[1]

ELVIN E. HERMAN, Manager, Advanced Development, Research and Development Laboratories, Hughes Aircraft Company, Culver City, California.

Factors presently inhibiting the acceptance of human engineering contributions, future areas of work, and means of expanding the utilization of human engineering in the electronics field are discussed from an engineering management viewpoint.

INTRODUCTION

In discussing the future roles of human factors in electronics, three aspects of the subject will be examined. The first of these will explore some influences which have considerably inhibited full human engineering utilization. The second aspect of this subject will highlight representative future areas of electronics in which human factors will warrant special effort. Lastly, recommendations will be advanced on possible ways to improve and to expand human factors utilization in the electronics field both in current type work and in new areas.

Turning first to the problem of human engineering acceptance and utilization by associated professions, there are several factors which appear to play an important part here. It should be kept in mind that these may or may not be a fault of the human engineering profession, but nevertheless they should be recognized. One of the most frustrating problems is that of belated input. Sometimes a particular project head, or in some cases a whole organization, fails to recognize the importance of human factors until after a design is in the field and it is proven unoperable. Then the human engineer is called in to convert a bastard system into usable gear without significant change — something we have heard many times. The restrictions here are obvious. Extreme compromises must be made to keep the hardware changes to a minimum. Inevitably additional design effort is incurred, the engineers' mistakes are highlighted, and the final result is a mediocre solution for everyone concerned. Even when human engineering work is instigated at program beginning, it is difficult to insure that research results are available in time to influence the original design.

A second factor which is oftentimes claimed as a reason for unacceptance of human factors engineers is that they lack operational experience. These complaints generally show up in two forms. For example, pilots claim (and perhaps rightfully so) that the real environment is much different than the simulation environment and the human engineer often does not have the operational experience to deny or confirm this fact. It therefore is questioned whether he is in a position to make the decisions with which he is entrusted. The other form of criticism is a resistance by the design engineer who claims that the human factors engineer does not recognize the complexities of the outputs or controls which he desires. Dr. George Long commented recently in a panel discussion that his own subjective experience gained in personally operating systems was not of great value to him in human factors engineering. It is very likely, however, that his influences are much more readily accepted because he is recognized for having this extensive experience.

A third factor which still inhibits to some extent the acceptance of human engineering by associated professions is simply the newness of the role, particularly in the fields of mission analysis,

[1] Presentation given at The Seventh Annual Human Engineering Conference, sponsored by the Office of Naval Research, 15–16 September 1959, Hughes Aircraft Company.

decision functions, and in preliminary equipment design. Here, the human factors engineer may be somewhat at fault for having overspent attention on relatively low-payoff items. This has created a prejudice in some electronic engineers' minds, so that their total conception of the role of the human factors engineer is to establish shape-coded knobs and optimum panel arrangements.

A fourth factor which contributes to the problems of human engineering is the delegation of human-engineering-type responsibilities to organizational units which are made up of personnel who do not have appropriate training.

There is a rapidly growing recognition of the importance of human engineering in military system programs, and in many cases specific human engineering studies, analyses, mock-ups, etc. are contractual requirements. The difficulty is that the authority is not necessarily delegated to qualified individuals. Instead of appropriately staffing organizational units and delegating corresponding authority, existing design engineering groups are merely renamed and a few human factors engineers are called upon to write reports justifying the human engineering decisions made by electronic engineers, a task that is sometimes not easy.

These then are four of the factors which are regarded as having inhibited the rapid acceptance and full utilization of human engineering personnel — namely: (1) the fact that they frequently are not brought into new programs early enough to make large and effective contributions, (2) the frequent and sometimes valid criticism that they lack sufficient operational experience and knowledge of engineering costs, (3) the newness of their roles in areas where they can make their greatest contributions such as mission analyses, decision functions, and preliminary equipment design, and (4) lack of delegated authority to those appropriately trained.

In general, the seriousness of these four factors seems to be gradually diminishing with time.

In any new application of a field of specialization, its rate of acceptance is influenced by many factors. These include mutual recognition of problem areas, importance of the new problem areas relative to the total program, the degree of education and to some extent the breadth of education of the personnel, the magnitude and quality of the contribution, and the maturity of the individuals. These are factors that should be concentrated upon by all participants irrespective of their field of specialization.

Having discussed some of the inhibiting factors, let us turn now and attempt to extrapolate a few future areas of emphasis for the human factors engineer associated with the electronics field.

One principal area which has received relatively little emphasis, except in a few organizations, is that of mission analysis from the human factors standpoint. Typically, mission analysis is carried out to establish requirements on subsystem and/or equipment designs, but it is also important to carry out this analysis from a human factors standpoint to establish the divisions of functions between automatic equipment and the man and, in cases where there is more than one man involved, between operators. These comments apply not only to manned vehicles, but also to the preparatory, checkout, and launching phases of unmanned devices as well as to earthbound systems. Mission analysis from the human factors standpoint is also required to establish operator decisions, the information requirements necessary for such decisions, the required accuracies of this information, and the displays and controls by which these decisions can be entered back into the system. Mission analysis from the human factors standpoint is also important to insure total mission integrity. In other words, to be sure that it is within human capability to carry out the tasks at the rates required, with the necessary accuracies, and in the sequences anticipated.

The second general area which foretells of greater human factors influence on electronics is in the research and engineering on operator

effectivity and handling capacity. One specific area is the determination of the relationships between input data resolution and decision rates, evaluation of the effects of environment, task loading, etc., on decision capacity, and finally the relative merits of various display configurations and scale factors. Here, it would appear necessary to evolve analytical models somewhat akin to those now in being in radar detection theory.

It might be well here to elaborate with a specific example. In the military concept of manned strike reconnaissance, the operator is called upon to examine map-type data at an extremely rapid rate and to effect identification in time to take appropriate actions. The equipment designer needs to know the resolution, dynamic range, and map scale or "window size" to provide because these considerations vitally influence equipment complexity. Factors critical to target identification must be determined and means of unambiguous target selection must be provided. In fact, the validity of the whole strike reconnaissance mission concept probably rests, to a major extent, upon the operator handling and identification capacity.

A third and very important area in which human factors will undoubtedly play an increasing role with electronics is in system maintenance, automatic checkout, and self-test philosophy and design. One particular area which has received insufficient attention is that of developing new training aids, i.e., means by which maintenance personnel can be familiarized with the equipment principles, trouble-shooting procedures, and general maintenance. Perhaps one of the most practical tools to accomplish such goals is the use of teaching machines. It is certainly evident that the teaching machine design itself is not the primary problem, instead the important area is in the appropriate programming of such devices.

The design of maintenance facilities and maintenance equipment itself is no less important than that of the basic systems, and here again the human factors engineer must share in the responsibility for the establishment of a maintenance philosophy and in generating requirements on the maintenance equipment as dictated by operator decisions. As with the primary system design, the human factors engineer must "follow through" to insure adequate maintenance equipment displays, readouts, and controls consistent with effective utilization, the maintenance philosophy, and the physical environment. In the maintenance area, it seems mandatory that the human factors engineer become personally familiar with the representative maintenance environmental conditions and procedures that are encountered by field personnel and not attempt to design without first-hand familiarity through personal observation.

Another general area which warrants increasing emphasis from the human factors standpoint is in system developmental and shakedown test planning. Most of our electronic design engineers have relatively little experience and only slight training in the design of statistical experiments and interpretation thereof. Consequently, many of our expensive test programs leave a great deal to be desired in the way of valid conclusions. Every branch of our military services today could profit by the careful application of statistical testing procedures and system evaluation; and although it is not necessarily a direct application of human factors, most human engineering personnel do have this vital element of training which is generally lacking in the field of electronic engineering.

Having explored a few of the new potential areas of human factors emphasis, let us reflect now on the problem areas mentioned earlier and suggest recommendations which would hopefully provide better qualified human engineers and will tend toward their more effective utilization in future roles.

The human engineering profession should take positive steps to expand the presently limited number of properly qualified human factors personnel and to improve the presently meager rate at which they are being produced by our universities.

In this regard several suggestions are in order.

In the long run more of our universities will have to organize the required interdisciplinary curricula and graduate research facilities. In order to do this effectively they should lure a few topnotch, experienced people back from industry and the government laboratories to guide them in the formulation of their programs and to get them started. To encourage the universities to do this and to help them provide students with the facilities and problems appropriate for graduate research experience, the government agencies and industry both should offer universities the assurance of the continuing availability of research contracts. The value of human factors research is certainly of sufficient importance to support a number of laboratories such as those at the University of Illinois and Ohio State during the decade following World War II.

To lend stability to the university programs, contracts should be let for a three-year period with one-year extensions each year. This is probably the most economical long-range solution to the shortage of human factors personnel.

A few words might be said about the kind of people that our universities should be providing. Most human engineering personnel appear more to be either psychologists with a smattering of engineering picked up principally through association, or else engineers with little or no formal training in psychology but with on-the-job experience with human factors problems.

There seem to be very few human engineering personnel with a genuine interdisciplinary background. Yet these are the kinds of people we will need in increasing numbers in the future. It isn't necessary that they all have PhDs. A five-year professional program culminating in a Master's degree could be designed to turn out people well qualified to practice human engineering in industry. These people would have a basic electronic engineering background including mathematics and the necessary physical sciences. Concurrently with this training — not just "tacked on" as an extra year — they would receive training in selected psychological subjects, including human engineering itself, physiological and experimental psychology, theoretical and applied statistics, and probability.

The PhD should of course be offered, and a person who wishes to specialize in research in human factors engineering would most likely follow this line. The point, however, is that there is a clear place for the professionally trained human engineer, a species that barely exists today since, to my knowledge, there are presently only two institutions in the country that offer such training.

It is easy to state that a professional curriculum in human factors engineering at our universities is desirable; it is quite another thing to bring such a curriculum into existence. However, the first step must be ours — in industry and in government. We must let our universities know our needs.

If a university did launch such a program, one of the problems would be that of finding qualified teachers for human engineering courses. Human engineering is in the interesting position of being an applied discipline yet without any large coherent body of theory to apply. Its practice therefore requires a considerable amount of practical experience. The teaching of human engineering, to be truly successful, must reflect this practical experience.

In order to provide this basic know-how to the young PhD who wishes to teach this subject, industry could be useful and at the same time help itself. Summer employment, an approach that we have tried here at Hughes, has proved to provide quite good results.

Turning now to the shorter range picture of providing qualified human engineering personnel, the following rapid-payoff procedures are recommended. These procedures involve a combination of on-the-job training and company-sponsored, part-time graduate study programs at local universities. This assumes that the company has its own experimental research facilities so that arrangements can be made with the universities for the employee-students to perform their research projects on real-life problems of

direct practical utility to the company. In our own experience some excellent Master's theses in biotechnology at UCLA have been produced in this way.

A second and eminently effective procedure in industry to provide the necessary broadening of training is for both the human factors personnel and the regular engineering personnel to take the time and the trouble to educate each other. For example, the human factors engineer should take the time to explain to the engineers and technicians the logic of the particular experimental design chosen, the experimental controls and the assumptions involved in the statistical analysis of the results. This will make the engineer much more sensitive in his own investigations to the adequacy of his data samples, the control of the independent variables and the like. Similarly the engineer can take the time to explain enough of the design details and approach for the human factors engineer to appreciate the perturbations that are possible. Only by the design engineer describing his approaches and design considerations can the human factors man learn to give proper consideration to such matters himself.

Although these cross-fertilization examples are obvious, over the years they pay handsome dividends to an engineering organization.

Turning away now from the subject of basic human factors training to the subject of more effective utilization in future roles, a few additional specific recommendations will be advanced.

Include in human factors organizations some personnel who have had system operating experience. Furthermore insure that human engineers have an increasing personal exposure to the operational, the maintenance, and developmental test environments; and they should be encouraged to personally participate in actual tests if at all feasible, or at least act as trained observers.

As discussed at length earlier, include in human factors organizations personnel who are dually trained in electronics as well as psychology, so that there is an intimate sharing of the design implication problems and the objectives, and so that there is a proper "bridge" between organizational functions.

Insofar as possible, insure that the *major* human engineering decisions are based upon firm objective scientific evidence rather than subjective personal opinions, but at the same time be sufficiently flexible and intuitive that an exhaustive quantitative simulation is not required for every *small-scale* decision.

Take special pains in human factors planning to insure that their primary efforts are devoted to those problems where the potential improvements to be gained are the greatest, and avoid those areas where the anticipated improvement is associated with a small percentage of the total task.

In mission analysis, preliminary system planning, and in early equipment design, involve human factors engineers directly with the engineering personnel so that they participate as a joint team in subsystem requirement planning and establishment of equipment logic, parameters, and configurations. This, to be really effective, must be done at the program beginning so that human factors personnel can perform a direct part in early system design rather than have to alter equipment after the fact. Then, as the design planning progresses, establish formal procedures for design reviews and here again insure that human factors personnel participate in these evaluations.

In conclusion, the field of electronics as well as the using organizations are demonstrating a growing appreciation for the value of human factors engineering. This role is an expanding one and, while many future military weapons will be unmanned, the complexity of man's tasks on the ground will require an ever-increasing attention by this profession.

A Comparison of Three Flight Attitude Displays

RAYMOND E. BERNBERG, Human Factors Group, Los Angeles Division, North American Aviation, Inc.

Requirements for presenting to the human operator the information needed to pilot air vehicles change as the performance characteristics of aircraft change. The most modern concepts of high altitude, supersonic air vehicles demand a more rigorous system for refined and accurate information for operator control with minimal error in performance than has previously been required. This study attempts to define the problem in a specific area of an optimal flight attitude display with which to fly a constant pitch attitude climb.

INTRODUCTION

There are special problems in the more advanced design supersonic air vehicles which require careful analysis of the ascent and climb-out to supersonic cruise. During the ascent and climb-out, small changes in pitch could prove excessive in fuel cost with consequent loss of range and time in flight. In fact, with the X-15, a one degree pitch error in flying a ballistic path could result in a 15,000 foot miss of the peak altitude needed to properly execute reentry. This problem creates new requirements in display information. Flight instruments with precision capabilities are needed in pitch attitude to minimize both loss of fuel and range.

Current air vehicles have utilized the attitude ball indicator coupled with reading the constant Mach and constant airspeed as indirect reading displays. However, in the case of a Mach three and Mach three plus supersonic air vehicles, the complex requirements for climb to supersonic cruise altitude demands a direct reading of a flight display with high precision in pitch attitude. During the segment of climb, no present instruments contain a constant readout with which to anchor a steady perceptual response for effective pitch feedback, inasmuch as flight instrument indications at this mission time segment are in motion.

PROBLEM

The logical question to ask concerning the capability requirement is, "Is the usual attitude ball indicator display adequate to accomplish the requirements of ascent and climb-out, or are there other flight attitude displays which might better accomplish what is required?" The latest WADC integrated flight display provides an attitude indicator with which to study this problem[1]. Essentially, the task involved for human performance is compensatory tracking. The problem revolves about the need to assess the latest tracking devices to compare human operator capabilities in meeting the system requirements.

METHOD

The X-15 flight simulator at Los Angeles was used to provide the experimental condition to compare flight instruments on the ascent to climb-out flight condition. The simulator contains a non-disturbed instrumented cockpit that has been integrated with a six degrees of freedom analog computer aerodynamic mechanization. It has complete cockpit presentation of controls and displays with which to evaluate flight characteristics. The X-15 is capable of an extreme high-angle flight and has among its flight instruments the WADC attitude indicator display containing an attitude ball indicator, a horizontal cross-hair bar, and a vernier scale side-read in vertical. This instrument is shown in Fig. 1. These three display features were selected for comparison under a simulated constant pitch attitude climb.

The attitude ball is marked grossly in ten-degree gradations. For pitch attitudes other than 0, 10, 20, and 30 degrees, the horizon bar must be read in an unmarked area of the ball

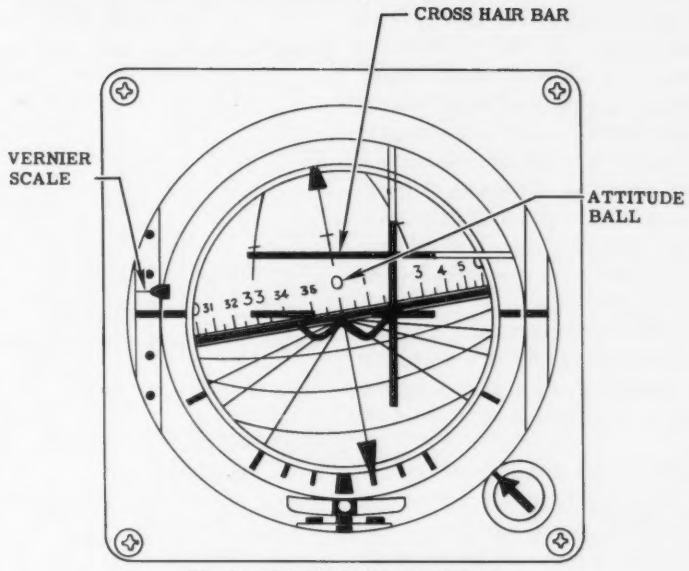

Fig. 1. Attitude indicator display.

for a command pitch angle. Within the area of reading for this study, the attitude ball has a 29 degree change for each inch of surface on the ball.

The cross-hair bar is capable of moving on command to a zero-read with the horizon. During a command pitch flight, the pitch angle is determined by an input to the scale, and therefore is the distance from the horizon to the angular distance away from the input. The cross-hair bar then must be moved through pitch change to zero-read with the horizon to achieve the command pitch angle. It also has the same relationship of angle change per inch as the attitude ball.

The vernier side-read is contained on the extreme left of the attitude indicator. It is a vertical scale with an arrow marker along one side that moves on command. During a command pitch flight, the pitch angle is determined by an input to the scale, and therefore is the distance from the central marker to the angular distance away from the input. The arrow marker then must be moved through pitch change to zero-read with the central marker to achieve the command pitch angle. Within the area of reading for this study, the vernier scale has a 9.5 degree change per inch of scale. Another variable was introduced by providing three levels of difficulty of flight control with differential damping through the stick to increments of aerodynamic stability. Fig. 1 illustrates the total display, in which is contained the three separate flight display conditions. The three levels were: 0 damping (no incremental change), 0.3 damping (small amount of sensitivity in stick action incremental to aerodynamic stability), and 0.6 damping (practically no sensitivity in stick action to aerodynamic stability).

PROCEDURE

The study was a full 3×3 factorial design, three conditions of stick damping and three instruments. All subjects acted in all conditions. Six North American Aviation, Inc., engineering test pilots were used to fly the experimental procedures.

The instructions to the subjects were to climb from a zero pitch angle to the command pitch of twenty-five degrees, and then hold this attitude for sixty seconds. The climb was to be a characteristically normal one which would reach ap-

proximately 2 G's and then level off to 1 G for the command pitch attitude. All pilots were to attempt to zero their roll at all times. Inspection of the flight characteristics of all subjects under the experimental conditions showed a close control of all the conditions required with no significant deviations. Each subject began his flight activity with a dry run for each instrument. He was then given the three different stick damping levels in separate order for one instrument before proceeding to the next instrument. The order of presentation of the three flight displays was different for each subject. In addition, the order for presentation of experimental conditions was equal in serial order for all subjects.

An oscillograph paper recording was made of flight activity for each subject. The sixty-second actual run indicated on the recording a continuous reading of adherence to the command pitch angle and the deviation from it. The data gathered was in units of angle change for each two-second period of the constant pitch run. Therefore, thirty samples of angle change for each subject were measured for each instrument with each level of stick damping.

RESULTS

Table I presents the means for deviation from command pitch under all the experimental conditions. Each unit of one is equal to ½-degree angle change per two-second period during the sixty-second flight run. The mean is computed from thirty samples of angle change for six subjects, or 180 measures per cell. The units were so devised as to eliminate the difficulty of handling small fractions in computation.

Table II presents the variance table with the sums of squares and resulting variance estimates.

Table III presents the F tests of significance, using both the interaction variance and the within-cells variance as a denominator. The interaction F is significant. This indicates that a source of variation exists in addition to the variables of stick damping, instruments and individual differences.

It is probable that the general trend of pitch angle deviation for instruments is affected in consecutive order by the differential levels of stick damping. See Fig. 2. The average pitch angle deviation using the vernier display is of the order of sixteen to one times as effective than using the attitude ball. The deviation from the twenty-five degree pitch command in the vernier display is less than 1/16th degree per two-second period of flight run.

The levels of stick damping have a significant F using the within-cells error term for the denominator. However, the F is not significant using the interaction error term. Inasmuch as the interaction F is significant, it is probably more appropriate to use this term to test the

TABLE I

Means for Deviation from Command Pitch N=6 (30 Readings* per Subject)

Stick Damping	*Instrument*		
	Ball	*Vernier*	*Cross Hair*
.6	2.216	0.155	1.950
.3	1.933	0.111	1.511
0	2.116	0.122	1.355

* One test unit equals ½ degree.

TABLE II

Analysis of Variance for Pitch Angle Deviation

Source	*DF*	*SS*	*Variance Estimate*
Stick Damping	2	22.38	11.19
Instruments	2	1125.12	562.56
Interaction Damping X Instr.	4	19.45	4.86
Individual Differences (Within-Cells)	1440	2521.81	1.75
Total	1619	3688.76	

TABLE III

F Test of Significance for Main Effects and Interaction

	Stick Damp.	*Instruments*	*Interaction*
Within Cells	6.39	321.46	2.77
P	<.001>.01	>.001	<.05>.01
Interaction	2.30	115.75	—
P	*>.05	<.001	—

* *Not significant.*

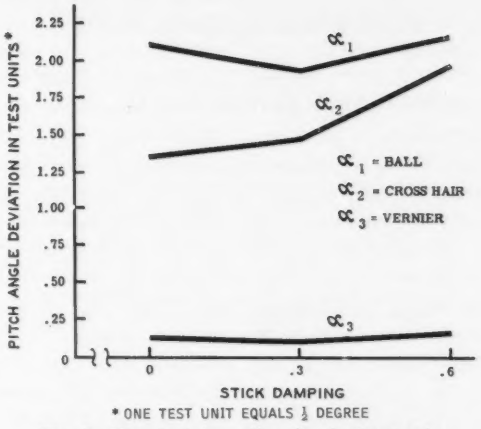

Fig. 2. Simple interaction for instruments with stick damping.

significance of the main effect. Fig. 2 again illustrates that the significance of differences between instruments are of greater order than any other factor in the study, thereby allowing pertinent generalizations, concerning the problems stipulated at the onset of the study.

DISCUSSION

It is quite apparent that the data demonstrates the inadequacy of the attitude ball to provide accurate information where a highly accurate command pitch is required. The cross-hair bar, although it provides a marker for zero-read, is also inadequate for the requirement under discussion. The grossness of sensitivity for both displays can only suggest time lost in flight, as well as fuel loss in an ascent and climb-out holding a constant pitch attitude. The vernier side-read scale appears to significantly demonstrate a minimal error display for this requirement.

It is interesting to note that the scale differences between the attitude ball and cross-hair bar with the vernier side-read are of a three to one ratio. The amount of improvement, however, is sixteen to one. The improvement is no doubt due to some relationship of scale factor and performance.

CONCLUSIONS

An experimental design, 3×3 full factorial, was effected with six engineering test pilots using the X-15 flight simulator. The conditions required the use of three flight displays: (1) vernier side-scale (2) attitude ball, and (3) cross-hair bar to fly a constant pitch climb of twenty-five degrees pitch attitude with three levels of stick damping. The results indicate the following:

1. The differences between instruments are highly significant.

2. The vernier scale is a more effective flight display under all conditions of the experiment than the other two displays. The vernier display is effective by a sixteen to one ratio over the attitude ball.

3. It is recommended that as a result of the evidence presented that consideration be given to the development of an attitude indicator display which will provide the base for optimum human performance, i.e., the use of a vernier scale in the attitude ball appropriately scaled for small pitch angle reading.

4. Further research is indicated in that a more precise knowledge of the relationship of scale factors with human performance is needed to establish more accurate servo-human dynamics relationships.

REFERENCES

1. SVIMONOFF, C., The Air Force Integrated Flight Instrument Panel, WADC Technical Report No. 58–431, WADC, WPAFB, Ohio; October 1958.

Speech Interference Level and Aircraft Acoustical Environment

JOHN J. DREHER and WILLIAM E. EVANS, Lockheed Aircraft Corporation, California Division, Burbank, California.

Basic concepts and formulation of the speech interference level (SIL) measure are discussed and the implications of the use of SIL to measure aircraft cabin environment are analyzed. Intelligibility tests with both words and phrases indicate that serious interference with speech can be demonstrated by adding supposedly unimportant frequencies to the SIL criterion masking band. Innocuous effects are also demonstrated by masking with low frequencies alone. Conclusions are drawn that because of the complex nature of a propeller-driven aircraft acoustic environment, SIL criteria, without qualification, are unacceptable for measuring either intelligibility or comfort.

INTRODUCTION

By mid-20th century we have accumulated a veritable mountain of data and information on the problem of voice intelligibility. We have had to learn, sometimes disastrously, that noise, whether it be electrical, acoustic, linguistic, or semantic, takes its toll on any kind of message intended for human reception. Acoustic noise, in particular, also takes its toll of human comfort and well-being, a fact of which the aircraft engineer is acutely aware. In recent years there seems to have been a tacit assumption on the part of engineers that if acoustic noise is capable of interfering with speech, it must *ipso facto* upset comfort in somewhat the same degree, and it is a matter of record that many such engineers presently use speech interference level (SIL) measures as a comfort yardstick in air vehicles.

The term "speech interference level" refers to the measure of the destructiveness of noise on the reception of speech, or "intelligibility." In its simplified form applied generally by the engineer, SIL has proved to be a useful method of comparing the relative effectiveness of speech transmissions under different environments of reception and is numerically defined as the arithmetical average of the sound pressure levels (in db re: .0002 dynes/cm^2) in the 600-1200 cps, 1200-2400 cps, and 2400-4800 cps octaves. This widely used simplification of the more cumbersome 20 band method of computing intelligibility developed earlier by French and Steinberg[2] is attractive from at least two points of view: its ease of use and its ease of understanding. However, where design responsibility is concerned, and where overdesign by a few decibels of sound control may represent needless expense and weight in an aircraft, it becomes apparent that the gross answer afforded by SIL becomes subject to scrutiny for possible refinement. The assumptions reviewed in this paper are:

(1) Speech intelligibility and comfort are synonymous.
(2) Different relative strengths of component SIL octaves are unimportant; i.e., only the average is significant.
(3) Contributions to speech interference by the noise spectrum outside the three designated octaves are unimportant for both comfort and intelligibility.

For an excellent discussion of the background and employment of SIL measures the reader is referred to Kryter[4].

Unfortunately, it is the low frequencies that most plague the aircraft engineer. In cabin environments of propeller-driven aircraft he will measure his greatest sound pressure levels in the neighborhood of "blade passage frequency." This is to say that the dominant input of noise at subsonic speeds is caused by the impulsive beats of air against the fuselage as churned up by the propeller blades. The frequency of this disturbance is equal to the number of blades times the rotational speed, thus, for a 4 bladed fan rotating 17 times per second, the blade passage frequency would be 68 cps, with considerable additional energy coming through at the

second harmonic, 136 cps. This situation holds true even though at speeds approaching 300 mph a high-frequency source, aerodynamic noise, begins to contribute a major share. It should be pointed out here that "aerodynamic noise" is not a noise field in the usual sense, but is the effect of the turbulent boundary layer flow pressure fluctuations exciting the fuselage in a manner equivalent to that of an externally generated noise — hence the term. Aerodynamic noise constitutes a broad band input (600 cps to past 10,000 cps), with a power that increases as a power function of indicated airspeed. Because, however, of the relatively high efficiency of insulating and trim materials in suppressing the medium and high frequencies, the low, blade passage frequency is still a chief offender at airspeeds of conventional propeller-driven aircraft.

We should review two things before describing the experiments which test the aforementioned assumptions, namely the SIL computation and the acoustic environment of the conventional aircraft cabin.

At present there are various ways of computing "speech interference level" although for our purposes it will suffice to note only the most popular, the Beranek, method[1]. This is the one most commonly used by aircraft engineers, and because of its familiarity enjoys a preferred place in the comparison of cabin noise environments. To arrive at the SIL value for any given broad band spectrum continuous in nature and lacking in impact noise intrusions (the necessary conditions imposed by the formulation) an octave band analysis is first made. Sound pressure levels in octaves 600–1200 cps, 1200–2400 cps, and 2400–4800 cps are measured in decibels re: .0002 dynes/cm² and arithmetically averaged to arrive at a single index. The higher this value, the lower the intelligibility score would be expected in the noise environment. Two things are thus implied: (1) that the frequencies below 600 cps are not important in determining the effects of noise on intelligibility, and (2) that with a given series of noises of varying broad band spectra a significant and high negative correlation should exist between intelligibility obtained in these noises and their SIL ratings.

The second desideratum, the acoustical environment of a typical propeller-driven aircraft spectrum, is shown in Fig. 1.

The ordinate, scaled in decibels re: .0002 dynes/cm² or db-Toepler (dbT)[7], indicates that the low-frequency step in the spectrum rides approximately 30 or 40 db above the "tail," which may be said to start at about 300 cps. In short, the strong negative slope of the curve gives evidence that, at the subsonic speeds maintained by all but the most advanced jet aircraft, the strongest reinforcement will be obtained in the region of blade-passage frequency.

Fig. 1 shows us the basic spectrum we have to deal with, and as the physical acoustician well knows, any reduction of the low end is dearly bought, since the low frequencies are relatively much harder to control than the middle and high frequencies. Any noise reduction effected in the cabin is likely to be done by suppressing, absorbing, or barring high-frequency energy, a condition that will result most probably in different SIL values even though the predominant character of the noise is set by the lows.

The experimental task before us now is to test intelligibility with masking noises of varying spectra and correlate the intelligibility scores with the SIL values of the masking agents. Most of the important frequencies for speech are usually conceded to fall within the span of 200 to 6000 cycles. The SIL rationale says, in effect, that we can predict how well speech will be received if we can specify the extent of interference in the 600–4800 cps region of the spectrum. We have seen earlier that ambient noise on the interior of an aircraft has, however, a very considerable low frequency component (i.e., below 300 cps).

The natural question arises: What effect, if any, does this strong, low frequency component have on intelligibility? There are two approaches we can follow: (1) test the interference of high and low components separately on the same verbal materials (the analytical

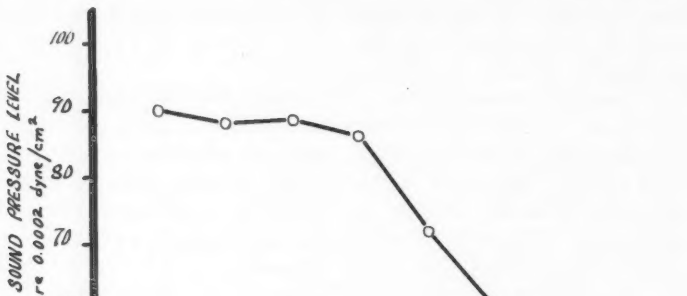

Fig. 1. Octave band analysis of typical turbo-prop transport aircraft cabin noise; normal cruise conditions.

method) and (2) compare these results to a combination of low- and high-component interference (the synthetic method).

STUDY I: MASKING EFFECTS OF SEPARATE LOW AND HIGH BANDS OF NOISE

Procedure

The first step undertaken was to test a given ensemble of speech signals under two different conditions: one using only "low-frequency" noise interference, the other using only "high-frequency" interference. The former consisted of filtered thermal noise from 50–300 cps, the latter of filtered thermal noise from 600–4800 cps (the SIL band frequency). Three types of verbal materials served to test intelligibility, PB (phonetically balanced) monosyllables, spondees, and 5-syllable phrases.

The PB words, representing the sounds of speech in approximately the balance afforded by the language, were the most difficult type of material of the three. The spondees, composed of two equally stressed syllables, were read beforehand by the subjects, and were written by the subject four times during the course of the test. The spondees were partially learned material, and their recognition would be analogous to that of a limited vocabulary of fairly high frequency of occurrence that might be employed in a standard radiotelephone operation, such as air traffic control. The phrases, composed of five contiguous syllables lifted from the context of naval flight instructions, preserved some of the redundance of continuous speech without being highly predictable because of syntax or subject matter. These would be analogous to a syntactical but unexpected message. Taken together, the three types represent a fair sample of the varied nature of spontaneous speech as it occurs in real life. The effects that are evident on such materials can properly be assumed to apply to speech communication in the general sense.

The 36 spondees were taken from lists W-1 and W-2, prepared by Central Institute for the Deaf, St. Louis[3], as were four sets of the phonetically balanced monosyllables (W-22 lists)[3]. The four sets of twelve 5-syllable phrases were taken from recorded instructions of naval flight instructors to their students[8].

To gather maximum data in a relatively short test, a "self-attenuating" format was employed.

In this method of testing the speaker's voice gets progressively fainter while the masking noise does not; in effect, then, several signal-to-noise ratios may be scanned in a relatively short time without the risk of aural overload induced by masking a constant voice level by increasing the noise.

First, an initial voice level was established by several seconds of "cold running" speech uttered at conversational rate and force (Overall Sound Pressure Level 70 db 1 meter from speaker). A broad-band noise was then recorded at the same average level on the leader of the test tape and its level considered as the index. Thus, speech played back at this level, masked by noise at this level, was considered a signal-to-noise ratio of 0 db. Speech 3 db below the index would result in a −3 db S/N ratio, and so on. Fig. 2 shows the recording setup. The microphone, an Altec 660A dynamic, was positioned at the corner of the speaker's mouth, the long axis lying along the cheek. This allowed precise control of the microphone-lip distance without introducing breath-blast.

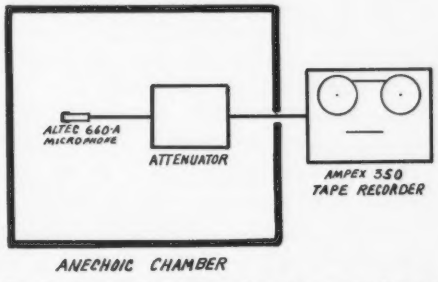

Fig. 2. Instrumentation plan for recording of test materials.

By an initial setting of the calculated average speech sound pressure level equal to that of the broad-band noise recorded on the leader of the test tape, subsequent progressive attenuations of the voice before the recording stage allowed presentation of all three kinds of verbal stimuli at each of four S/N ratios: −3, −6, −9, and −12 db. To achieve still lower S/N ratios without getting too near to the absolute auditory threshold, the same progression of attenuation was employed a second time, with a masking noise 9 db more intense, yielding effective ratios of −12, −15, −18, and −21 db. Such extreme relative attenuation was necessary because of the expected differential in masking capabilities of the two bands.

Subjects and Test Presentation

Eight subjects, engineers and technicians of the Physics Research Department of Lockheed Aircraft Corporation, were tested individually in the anechoic chamber. Both the stimuli and the masking noise were presented via loudspeaker, with the subject seated six feet distant from, and on the axis of, the two adjacent speaker units. Speech was delivered by a 5″ Utah speaker, of the type recommended for installation in Lockheed aircraft. Noise was furnished from a separate source by a 10″ high-fidelity AR-2 speaker system. Both speakers gave sufficiently good response for the purpose intended: the Utah response was relatively flat between 150 and 6000 cps; the AR-2 response was flat between 50 and 10,000 cps.

Masking noise was made by filtering the output of a General Radio Random Noise Generator through an Allison 2A Band-Pass Filter, and sending the filtered band through a PERI 50 watt amplifier to the AR-2 system. Speech was fed from the Ampex 350 recorder to a Gott G 30 U amplifier, thence to the Utah speaker. Fig. 3 shows this schematically.

Prior to testing, each subject was requested to study the list of 36 spondees and to read them aloud. He was informed that these, and only these, words would constitute the test lists of spondees, which would be indicated on his test blank. He was informed of the nature of the PB words and 5-syllable phrases, but cautioned that they would be different for every level and S/N ratio tested. He was further informed that, although the level of the masking noise would remain constant in any one test condition, the voice would get fainter and fainter, and that his job was to get all the words presented, guessing on those of which he was not sure.

After his questions were answered, the test

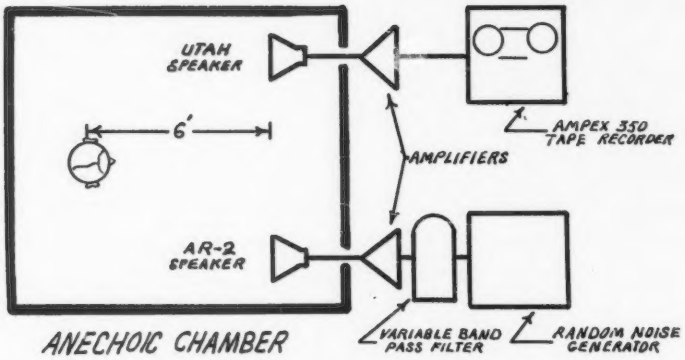

Fig. 3. Instrumentation plan for presentation of Study I.

commenced. The sequence consisted of four test blanks, each blank allowing space for 36 spondees, 12 PB's, and 12 five-syllable phrases. Masking noises were identified as follows:

TABLE I

Masking Noises for Single Band Test

Noise	*Over-all Level*	*Frequency, cps*
A_1	80 dbT	50–300
B_1	80 dbT	600–4800
A_2	89 dbT	50–300
B_2	89 dbT	600–4800

To avoid any systematic bias that might arise from standardized presentation order, the masking was scheduled $A_1B_1A_2B_2$ for half the listeners and $B_1A_1B_2A_2$ for the other half. In all cases subjects heard the lower (80 dbT) noise level first to avoid auditory fatigue that might possibly occur at the higher (89 dbT) level.

Results and Discussion

LOW Band: Fig. 4 shows the results of intelligibility scores for all three types of verbal materials. It will be observed that comparatively little masking effect is noted with the 50–300 cps noise band; in fact, only with the unfamiliar PB monosyllables is any effect of the LOW band seen, and this is most probably attributable to sampling variations in the words selected in this category.

SIL Band: The interference with speech caused by the 600–4800 cps masking is quite apparent from inspection of Fig. 4. In the process of going from −3 db to −21 db S/N ratio, the attrition was definite for all materials, with the most severe damage occurring, as expected, to the monosyllables.

Here, incidentally, is a good illustration of how conclusions about systems can be colored by the particular test materials. At −6 db, an undesirable but not uncommon S/N ratio in air communications, the results of the PB test would indicate a much lower efficiency than those of the spondee and phrase tests, yet the actual use of the system is much better represented by the spondees and phrases. Furthermore, from −9 db to the end of the range, −21 db, the PB's offer almost no gradations of difference and do not indicate the actual pattern of attrition to connected discourse as introduced by the noise. These considerations should be kept in mind when cost decisions are to be made on the basis of tests. This is simply to say that it takes a better system to get high scores on PB words than on syntactic material and that the PB words give a more exacting test.

One cannot, however, assume from these results that low-frequency masking may be disregarded, nor that a predictive formula may be based upon only the high-frequency components simply because in isolation only the latter seems to have any effect. The actual environment of

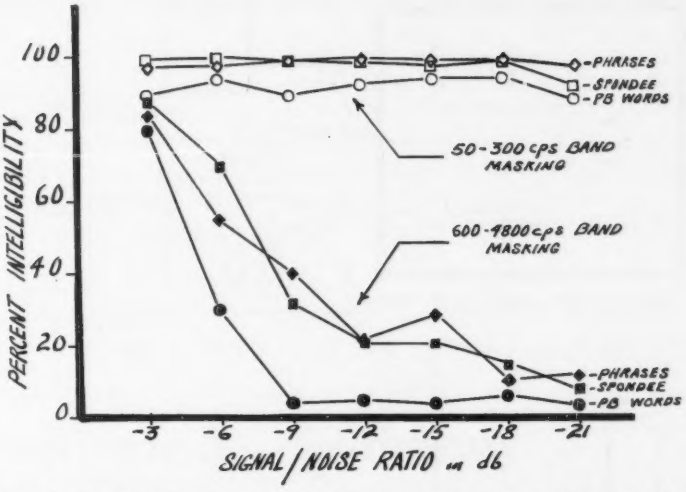

Fig. 4. Intelligibility scores for PB words, spondees, and phrases masked by low band (50–300 cps) and SIL Band (600–4800 cps) noise.

the aircraft passenger, as pointed out previously, contains both bands, with the lower band usually predominant in strength. To test this situation, a synthetic method must be used.

STUDY II: MASKING EFFECTS OF A COMBINED LOW- AND HIGH-FREQUENCY NOISE BAND

Procedure

The preceding test has indicated that virtually no upset to speech is occasioned by low-frequency masking alone over a wide range of S/N ratios. Intuitively we might suspect that any kind of noise, if intense enough, would destroy speech, and, indeed, this is seen to be the case with Pickett's[6] tests of low-band masking. Certainly the "typical" aircraft ambient cabin spectrum (see Fig. 1) might be suspected to involve some of this effect. The plan for this test, however, is to measure the effect on intelligibility by adding a LOW band (50–300 cps) of a given sound pressure level to a SIL band (600–4800 cps) of equal level to form a complex masking noise. If the addition of a low band at this relative magnitude causes any effect, it may be assumed that a relative increase of low band strength would serve only to intensify this effect.

Since one month had elapsed since the previous test had been performed, sufficient extinction of stimuli recognition was assumed to have occurred to allow use of the same test tape in the synthetic phase of this investigation. This procedure, when possible, is a very desirable practice, since use of the same stimuli represents the most rigid possible experimental control. The description of its preparation and makeup has been described above.

Because two bands of masking noise, each independently variable, had to be supplied, two noise generators were employed, with their filtered outputs adjusted for level and mixed in a multichannel amplifier prior to being led to the appropriate speaker. Again, a separate speaker system was used for the stimuli. The equipment and arrangement is identified in Fig. 5.

The speaker systems used were a Stevens FR 80 for speech, a Jim Lansing for noise. Speech, from an Ampex 600 tape recorder, was fed through a Holt Stereo amplifier. The random noise was filtered into two bands, the LOW by an Allison 2A Variable Filter, the SIL band through an MSA Soundscope filter section. Both

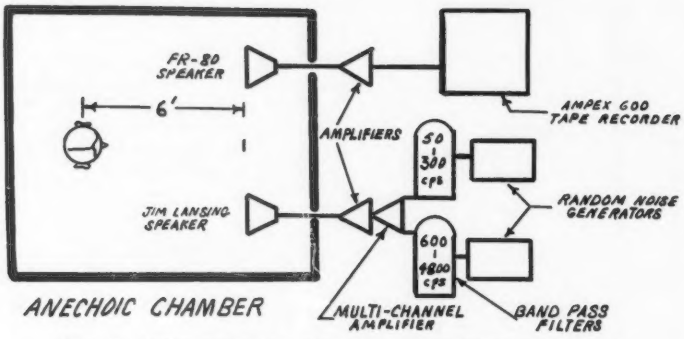

Fig. 5. Instrumentation plan for presentation of Study II.

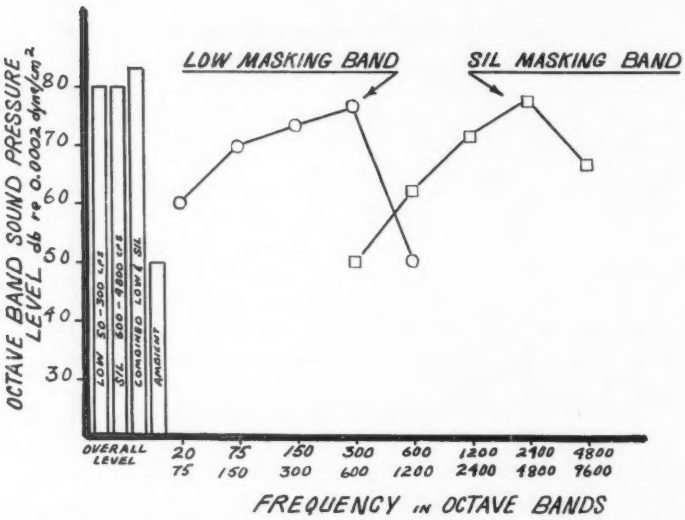

Fig. 6. Octave band analyses of low band (50–300 cps) and SIL band (600–4800 cps) masking noises.

bands were then synthesized in a specially constructed multichannel precision attenuator amplifier and led to a Gott G 30 U power amplifier. The profile of these masking bands is given in Fig. 6.

Subjects and Test Presentation

Subjects were tested individually in the anechoic chamber. Since the same subjects served in this phase, they were retrained on the spondees in the same manner as before, perusing the list and reading the words aloud. Their instructions on the mechanics of the test were the same as previously. The test in this phase consisted of four separate batteries of words and phrases. A complete battery presented both words and phrases at four diminishing S/N ratios, and was given under one masking noise. This was followed by another battery using both masking noises simultaneously, thus alternating masking conditions until the four batteries were completed (Order: ABAB). Alternate subjects, a total of half the panel, were given a second presentation order of BABA to effect a

counterbalance of fatigue and accommodation factors.

A burst of white noise on the leader of the tape was again used as a level-setting index. The recorded voice was attenuated in relation to this index to achieve the ratios of −3, −6, −9, and −12 db at the listener's position, and it was decided that lower ratios would offer no useful information in view of the results of the preceding test. Consequently, a masking level of 80 dbT was adopted only for the 600–4800 cps mask, and setting of the index to that level assured the four ratios. The complex (two-band) mask was composed of the SIL band at 80 dbT plus the LOW (50–300 cps) band fed in at equal sound pressure level. This resulted in a new base level of 83 dbT, against which the index of the speech level was set for this part of the test to maintain the given S/N ratios as before. A change in absolute level of this amount results in no effect on speech reception, since it is a function of the S/N ratio within rather wide limits.

Fig. 6 will give the reader some idea of the nature of the two noises, the complex mask being a two-humped curve with peaks of equal height.

RESULTS AND DISCUSSION

Intelligibility scores as a function of the S/N ratio and type of masking may be seen in Fig. 7.

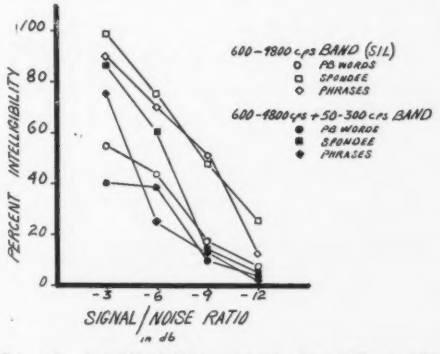

Fig. 7. Intelligibility scores for PB words, spondees, and phrases masked by SIL band (600–4800 cps) and complex band (50–300 cps + 600–4800 cps) noise.

The lower curve in each graph, representing speech masked by the complex noise, shows the effect of adding what proved to be, when tested in isolation, masking of no importance. The solid line, connecting the values obtained when using only the SIL band, is similar to the results obtained previously. Fig. 8 collapses scores over all S/N ratios and gives a picture of the general attrition experienced with each kind of masking.

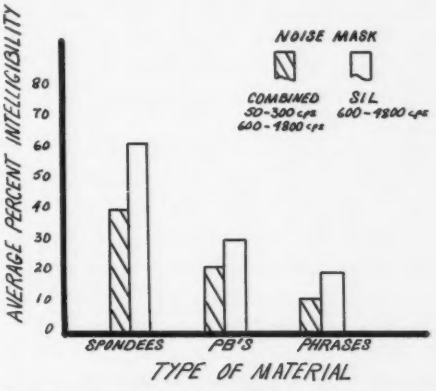

Fig. 8. Average intelligibility scores taken over all signal noise ratios tested for PB words, spondees, and phrases masked by SIL band and complex band noises.

There is something more to be learned from the test data than a mere confirmation that addition of low-frequency noise components decreases intelligibility. The construction of the test allows examination of the particular nature of the destructiveness. Two kinds of damage to speech can result from noise interference: (1) the signal may simply be blotted out so that the listener cannot perceive it, or (2) the interference may be of such a nature that the message will be confused with other possible messages. Of the two error types, the latter is more dangerous, since the receiver may assume he has heard correctly and act on the spurious message. If he knows he has missed the signal, he may always ask for a repeat transmission. If he suspects only that he might have

misheard, the probability of asking for a repeat is lessened.

Statistical linguistic theory would imply that some accommodation for the listener could occur under confusion type masking (i.e., that containing low frequencies). If the message contains a relatively large number of acoustic and syntactic cues, as is the case with connected discourse or longer words, the greater will be his probability of being able to utilize these by practice in listening or by repeat transmission. If, however, the character of the noise is such that the signal is actually "destroyed," our listener has reached a perceptual "floor" and can be helped only by mechanical betterment of transmission.

Inasmuch as four batteries of tests were used here, two batteries, one for each type of masking, were, in essence, retests. This is to say, when listeners, in the course of the experiment, were subjected to four batteries but only two kinds of masking, both masking agents were tested twice, and the second time they heard test words under, let us say, the single-band mask, they would generate a score that might differ from that achieved during their first exposure. This allows us to examine both efforts and determine whether any statistically significant improvement was made on the second hearing. Where the acoustic cues are sparse to begin with, as in the case of the PB words, the chance exists that no amount of repetition of the test conditions would allow any improvement. Where, however, either additional acoustic cues (as with the spondees) or additional syntactic cues (as with the phrases) exist, the possibility of increasingly efficient utilization of these is present. The best available evidence of this lies in a comparison of the group intelligibility scores on the primary and secondary batteries for each kind of material, and is referred to as an "intra-trial" by comparison. By collapsing the scores made on both trials for each of the two masks, we can also compare the effects of the masking destructiveness. This is referred to as an "inter-trial" comparison, and from the display of values in Fig. 8 would be expected to be significant. Tables II and III, a treatment of these data by Student's *t*, gives intra-trial assessment for the single- and double-band noises. Table IV gives inter-trial comparisons between the two masking conditions. All scores are collapsed over four S/N ratios.

TABLE II

Comparison of Intra-Trial Intelligibility Values of Three Types of Verbal Material Tested Under Single-Band (SIL) Noise Masking

	Trial						
Material	*1*	*2*	*Diff.*	*t*	*df*	*p level*	*Sig. Improvement*
Spondees	58.6	62.7	4.1	2.9	8	.02	Yes
PB's	25.9	25.0	−0.9	1.1	8	.30	No
Phrases	54.5	52.0	−2.5	2.4	8	.05	Yes

TABLE III

Comparison of Intra-Trial Intelligibility Values of Three Types of Verbal Material Tested Under Double-Band (SIL plus LOW) Noise Masking

	Trial						
Material	*1*	*2*	*Diff.*	*t*	*df*	*p level*	*Sig. Improvement*
Spondees	32.0	40.4	8.4	3.1	8	.02	Yes
PB's	20.4	22.2	1.8	2.7	8	.05	Yes
Phrases	25.7	31.8	6.1	8.4	8	.001	Yes

TABLE IV

Comparison of Intra-Trial Intelligibility Values of Three Types of Verbal Material Tested Under Two Types of Masking Noise

	Mask						
Material	*SIL*	*Mixed*	*Diff.*	*t*	*df*	*p level*	*Sig. Diff.*
Spondees	60.7	38.9	21.8	3.2	8	.02	Yes
PB's	25.4	21.3	4.1	2.7	8	.05	Yes
Phrases	53.2	29.4	23.8	3.1	8	.02	Yes

Table II, comparing the improvements due to repeated trials during the course of the test, shows that only the monosyllables were insensible to some sort of learning effect, and this only under SIL band masking alone. In all other test replications a significant improvement was noted with all types of verbal materials. Table IV, contrasting the results of SIL band masking and SIL-plus-LOW masking, shows the exceedingly heavy toll taken on test scores by adding to the SIL band the apparently innocuous 50–300 cps noise energy. It would thus appear that with spectra of propeller-driven aircraft type, with strong low components, speech interference is not accurately forecast by the SIL band computation, and that sound control measures affecting the low frequencies could not, therefore, be reflected by the SIL method.

Additional work[5] at Lockheed Aircraft Corporation to be reported later shows that although the destructive effects of noise on speech decrease from the beginning to the end of the 600–4800 cps (SIL) range, precisely the opposite effect is true with subjective annoyance reaction. That is, the annoyance mounts steadily as the frequency of the interfering noise rises, giving a negative correlation with intelligibility. This also negates the implication that different SIL octave strengths resulting in the same average level yield the same annoyance magnitudes. According to these researches a single average SIL value could result in a wide spectrum of annoyance responses.

CONCLUSIONS

In this paper the use of the SIL value as a measure of acoustic comfort was questioned as to validity. The experimental evidence indicates that under conditions of broad band noise with strong low-frequency reinforcement, the low noise components materially affect intelligibility even though the same low components acting alone affect intelligibility little if at all.

It is also shown that syntactic verbal materials are more seriously deranged than are monosyllabic test materials.

It is concluded that SIL is unsatisfactory as an index of aircraft cabin noise environment.

REFERENCES

1. BERANEK, L. L., Noise Control in Office and Factory Spaces. *Trans. Chem. Eng. Conferences,* Industrial Hygiene Foundation of America, Nov. 15, 1950.

2. FRENCH, N. R., and STEINBERG, J. C. Factors Governing the Intelligibility of Speech Sounds. *J. Acoust. Soc. Amer.,* **19**, 90, 1947.

3. HIRSH, I. J., The Measurement of Hearing, pp. 328–333, 1952. McGraw-Hill: New York.

4. KRYTER, K. D., Speech Communication in Noise. Air Force Cambridge Research Center TN 54–52, May, 1955.

5. Lockheed Aircraft Corporation. Studies in Annoyance Effects of Noise–II, Physics Research Memorandum 8111, Burbank, Calif. Oct., 1958.

6. PICKETT, J. M., Low Frequency Noise and Speech Intelligibility. *J. Acoust. Soc. Amer.,* **31**, 1259, 1959.

7. THORPE, HOWARD A., On Abbreviating db re 0.0002 Microbar. *J. Acoust. Soc. Amer.,* **26**, 871, Sept., 1954.

8. WALKER, C. W., and BLACK, J. W., The Intrinsic Intensity of Oral Phrases. Report NM 001 064.01.02 U.S. Naval School of Aviation Medicine, Pensacola, Fla. May, 1950.

Optimum Symbols for Radar Displays[1]

HUGH M. BOWEN, JOHN L. ANDREASSI, SHAFFER TRUAX, and JESSE ORLANSKY, Dunlap and Associates, Inc., Stamford, Connecticut.

Experiments were conducted to determine: 1) sets of geometric symbols which can be discriminated from each other and recognized with high accuracy under a variety of display conditions, especially those involving degradation of the image in ways similar to those that occur on radar displays, and 2) the size and strokewidth to height ratio desirable for symbols to be used on complex displays.

The recommended symbols are illustrated in the report; symbol height should be 1/2 inch or more and the strokewidth should be 1/8 to 1/10 of the height.

Suggestions for combining auxiliary symbols with primary symbols are given.

STATEMENT OF THE PROBLEM

Many military and civil operations require descriptive information to be displayed, generally in the form of geometric symbols. One step in the establishment of a standard symbol code is the measurement of the recognition value of different symbols. A successful code will be one in which observers are able to recognize each symbol with minimum confusion even under adverse conditions of display.

BACKGROUND TO STUDY

Previous work has indicated that certain shapes are more distinctive than others[2, 3], but it is difficult to establish a set of construction rules which will provide recognizable shapes[1, 4]. Recognizable shapes generally conform to some or all of the following requirements.

- Simplicity
- Symmetry
- Continuous contour (and often long perimeter)
- Relatively large enclosed area
- Either sharply angular or simple curved forms
- Familiarity — especially in the sense of having a familiar name and/or meaning.

Even when these conditions are met, to predict the recognition value of a particular symbol, the other symbols from which it must be distinguished have to be considered. For example, a rectangle is recognized readily from the set: rectangle, circle, cross; but is likely to be confused in the set: rectangle, square, pentagon. In addition, the visual size of the symbol itself and its various parts affect its recognition value.

These considerations led to two experiments. The first experiment measured the probability of identifying each of a set of 20 symbols when they were observed under a range of degradation conditions; the second experiment determined the size of symbol and the width of its strokes which would make it easily visible.

STUDY I. RECOGNITION ACCURACY OF 20 SYMBOLS

Experimental Conditions

Seven young male observers viewed $\frac{1}{2}$ inch high symbols (light on a dark ground) which were projected one at a time for 0.5 seconds on a screen. They were viewed at a distance of 50 inches at a brightness level of 4 to 5 foot lamberts. The 20 symbols are shown in Fig. 1. The various degradation conditions (Fig. 2) imposed were:

1. Obscuration by "noise"— none, moderate, or severe.
2. Distortion — not distorted or distorted.
3. Blur — not blurred or blurred.

The observers were provided with an answer sheet showing a clear picture of each of the 20

[1] This work was carried out for the Office of Naval Research, Washington, D.C., under Contract No. NONR 2682(00).

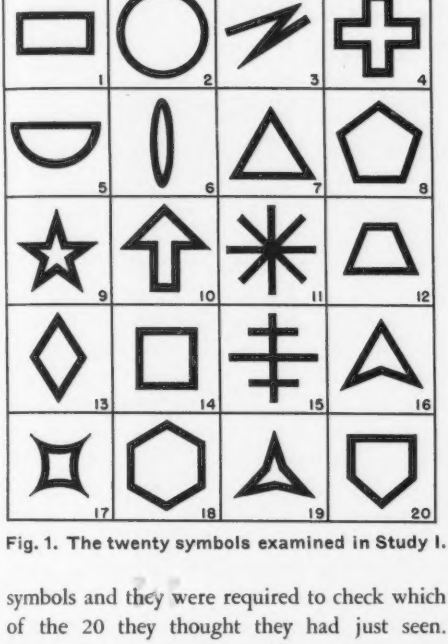

Fig. 1. The twenty symbols examined in Study I.

symbols and they were required to check which of the 20 they thought they had just seen. Hence, the experiment tested how well the various symbols under varying conditions of display were recognized when viewed separately, and when the observers knew the set of 20 symbols which could appear.

Results

The symbols can be recognized in the same order as shown in Fig. 1, number one being the easiest to recognize. It is recommended that only the first ten symbols be used. Table I gives optimum sets of symbols.

Under adverse conditions of display use only symbols 1 through 6.

STUDY II. SYMBOL SIZE AND STROKEWIDTH

Experimental Conditions

Three symbol sizes were used, 0.25″, or 0.375″, or 0.50″ squared. For each size of symbol three different strokewidths were used:

TABLE I

Optimum Sets of Symbols. (Symbol numbers refer to identification numbers used in Fig. 1)

Number of Symbols in Set	*Recommended Symbols*
2	1 & 2; or 1 & 3; or 2 & 3; or 7 & 14; or 5 & 7; or 5 & 14.
3	1, 2, & 3; or 5, 7, & 14.
4	1, 2, 3, & 4; or 5, 6, 7, & 14.
5	1, 2, 3, 4 & 5; or 4, 5, 6, 7, & 14.
6	1, 2, 3, 4, 5, & 6.
7	1, 2, 3, 4, 5, 6, & 7.
8	1, 2, 3, 4, 5, 6, 7, & 8.
9	1, 2, 3, 4, 5, 6, 7, 8, & 9.
10	1, 2, 3, 4, 5, 6, 7, 8, 9, & 10.

strokewidth to height ratios were 1:6, or 1:8, or 1:10. Four symbols were used: a circle, a triangle, a square and a variation of a cross (symbols numbered 2, 7, 14 and 15 in Fig. 1).

For each of the 9 conditions of symbol size and strokewidth two displays were made. Each display consisted of 40 white symbols distributed evenly on a 12″ display. Considering the two displays together, each symbol appeared 20 times, but on any one display it could appear from 7 to 13 times.

Eight young male observers were used, and their task was to count the number of times a specified symbol appeared on a display as accurately and quickly as they could. Viewing distance was 7 feet; brightness was approximately 20 foot lamberts. In the instructions, accuracy was emphasized over speed. Error and time scores were obtained.

Results

The results appear in Table II where the entries are the average time in seconds to count one symbol. Few errors were made. Where errors were made a small increment was added to the time score in order to make it comparable to the majority of cases where no error was made.

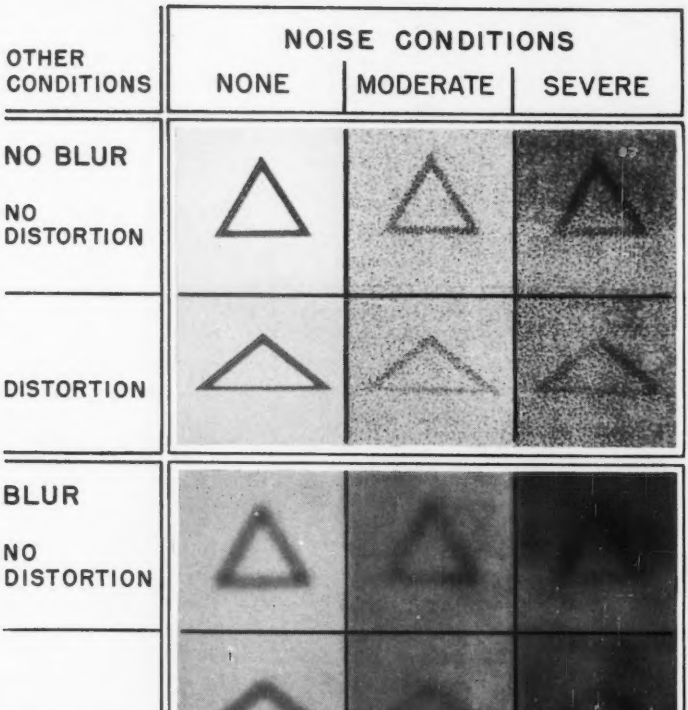

Fig. 2. Degradation conditions of Study I.

TABLE II

Average Time to Count Symbols of Varying Size and Strokewidth.

Size in inches	*Strokewidth/Height Ratio* 1:10	1:8	1:6	*Average for all strokewidths*
.25″	.93 sec.	.96 sec.	.99 sec.	.96 sec.
.375″	.74	.74	.80	.76
.50″	.71	.70	.70	.70
Average for all sizes	.79	.80	.83	

The 0.50 inch symbols were counted fastest; for this size of symbol the strokewidth to height ratio makes no difference. When the symbols were smaller, thinner strokewidths gave better performance.

It is recommended for near viewing distances (up to 10 feet) symbols should be 0.4″ or more in size (i.e., fit into a square 0.4 × 0.4 inches). Strokewidth to height ratios should be between 1:8 and 1:10 (i.e., strokewidths of 0.04″ to 0.08″). Where the viewing distance is longer the symbols should be large enough to make an angle at the eye of 22 minutes or more.

The average times taken to count each of the symbols were:

Cross	0.69 seconds
Circle	0.77 "
Square	0.84 "
Triangle	0.93 "

In operational situations where the operator must appreciate the numbers of various symbols

that are present, one second per symbol is most likely the fastest, practical performance that can be realized.

CONCLUSIONS

Symbols should be chosen from the sets given in Table I. Extra symbols should be avoided. When the display degrades the legibility of symbols every effort should be made to keep the symbols few in number and preferably chosen from the best six symbols.

Symbols should be about ½" high or more and have a strokewidth ⅛th to ⅒th of their height. If small symbols must be used, the strokewidth to height ratio must be 1:10 or less, but with the proviso that the strokewidth should not be smaller than 0.02 inches. When display conditions are poor, symbols should be increased in size.

AUXILIARY SYMBOLS

It is often required that primary symbols have added onto them auxiliary symbols, somewhat in the manner of the meteorological station code. While no experiments were conducted on this problem, there was an opportunity to observe the use of present symbol codes.

The following precepts suggest the manner in which complex codes should be built:

1. Primary symbols should be large and enclose a space.
2. No auxiliary symbol should cross, distort, interfere with or in any way obscure the primary symbol(s).
3. Symbol complexes should not normally exceed two geometric symbols (possibly three in some circumstances), a location dot, and a speed and direction vector line where applicable.
4. When other information is required it should be in the form of numerical representation (e.g., one, two, or three marks to indicate magnitude of object) or in actual numbers and letters.
5. The geometric center of the symbol and/or a large clear dot should indicate location.
6. Auxiliary marks should be compact solid figures. Fig. 3 illustrates some of the above points.

Symbol codes can quickly lose their advantages and efficiency if too much information is portrayed. Hence, coded information should be kept to necessary essentials. An effective device for handling a great quantity of information is to place certain categories of information on a selective basis, so that they do not usually appear on the scope but can appear "on demand."

REFERENCES

1. ATTNEAVE, F., Physical determinants of the judged complexity of shapes. *J. exp. Psychol.*, **53**, 221–227, 1957.

2. CASPERSON, R. C., The visual discrimination of geometric forms. *J. exp. Psychol.*, **40**, 668–681, 1950.

3. FRENCH, R. S., Pattern recognition in the presence of visual noise. *J. exp. Psychol.*, **47**, 27–31, 1954.

4. WULFECK, J. W., and TAYLOR, J. (eds.), *Form Discrimination as Related to Military Problems* (Proceedings of a symposium sponsored by the Armed Forces — NRC Committee on Vision at Tufts University, Medford, Mass., April, 1957) National Academy of Sciences, National Research Council, Washington, D.C. Publication 561, 1957.

SUPPLEMENTARY MATERIAL TO REPORT

Recognition Accuracy of 20 Symbols

The study determined the accuracy with which observers could recognize twenty different geometric symbols, under various degrees of degradation.

The experimental plan was a complete five-way factorial with single replication.

Subjects	7	
Symbols	20	
Obscuration	3	(no noise, moderate noise, severe noise)
Distortion	2	(not distorted, distorted)
Blur	2	(not blurred, blurred)
Replications	2	

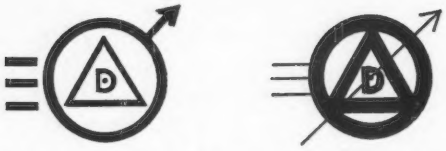

Fig. 3. Good and bad practice in symbolic code construction.

For the purpose of analysis of variance each response to each symbol presented was scored correct or incorrect. Entry into the analysis of variance matrix was for incorrect responses. The complete analysis of variance table is given in Table I.

The p values given in the right-hand column of Table I are those which were used for interpreting the results of the study. In all cases they were obtained by using as an error term the highest value and significant term in the immediately superior interaction order which contained the term or terms whose variance was to be tested. This conservative procedure was followed as it was felt important to avoid any Type I errors in the interpretation of the data.

A confusion matrix analysis was performed on the data and Table II gives the probability (average over all display conditions and for the seven subjects) of any symbol being responded when any certain symbol was presented. Hence each cell entry is based on 168 observations. The numbers in italics along the edges are the symbol numbers used in the report. In order to find optimal subsets of symbols the Constant Ratio rule as developed and validated by Clarke was used (Clarke, F. Constant Ratio rule for confusion matrices in speech communication, *J. Acoustical Soc. Amer.*, **29**, 715–720, 1957).

TABLE I

Source of Variance	*d.f.'s*	*Sums of Squares*	*Variance Estimate*	*p*
Subjects (S)	6	6.64	1.11	
Symbols (B)	19	65.26	3.43	$<.05$
Noise (N)	2	52.49	26.25	$<.05$
Focus (F)	1	11.79	11.79	
Distortion (D)	1	54.27	54.27	$<.01$
S×B	114	22.24	.20	
S×N	12	5.68	.47	
S×F	6	1.70	.28	
S×D	6	3.12	.52	$<.01$
B×N	38	15.84	.42	
B×F	19	4.57	.24	
B×D	19	26.85	1.41	$<.05$
N×F	2	1.00	.50	
N×D	2	4.42	2.21	
F×D	1	.07	.07	
S×B×N	228	23.09	.10	
S×B×F	114	11.67	.10	
S×B×D	114	16.82	.15	$<.01$
S×N×F	12	3.85	.32	
S×N×D	12	4.04	.34	
B×N×F	38	4.06	.11	
B×N×D	38	23.10	.61	$<.01$
F×D×N	2	2.19	1.10	$<.05$
F×D×B	19	3.52	.19	
F×D×S	6	1.37	.23	
S×B×N×F	228	26.01	.11	$<.05$
S×B×N×D	228	20.04	.09	
S×B×F×D	114	11.76	.10	
S×N×F×D	12	2.89	.24	$<.01$
B×N×F×D	38	6.78	.18	$<.01$
S×B×N×F×D	228	21.46	.0941	
Within	1680	152.50	.0907	
Total	3359	611.09		

TABLE II

OUT		IN																		
		4 1	9 2	11 3	7 4	12 5	15 6	16 7	14 8	17 9	1 10	10 11	18 12	13 13	5 14	8 15	6 16	19 17	2 18	20 19	3 20
4	1	.869				.006	.042		.018	.006	.006	.042	.012	.006							.006
9	2	.030	.839	.030	.024	.018	.024	.018		.030	.006	.012	.006	.024	.012		.012	.024	.006	.006	.012
11	3	.018	.012	.785	.006		.048		.006	.036		.006			.012	.012	.012	.018		.012	.036
7	4	.006		.006	.875	.048	.018	.149		.006			.006	.024				.208	.012	.006	
12	5	.006	.006	.012	.018	.756	.006	.006	.006	.060	.006	.006	.018		.024	.030		.012	.012		
15	6	.012	.012	.036			.762		.006	.018	.012	.012		.006	.012	.006	.012	.018		.006	.018
16	7	.006			.006	.006	.006	.553		.006						.018		.065		.006	
14	8			.006					.506	.113			.006	.012		.012	.006	.006	.006	.018	
17	9		.042	.006	.006	.012	.006	.030	.006	.458	.018	.012		.018		.012	.012	.018	.006	.006	.006
1	10			.006		.012	.006	.012	.351	.048	.916		.024	.006		.012	.006	.006		.024	
10	11		.024		.018	.006	.006		.006	.012	.012	.863		.006		.006	.018	.012		.006	.018
18	12	.006	.006		.006	.030		.006	.012	.018		.006	.720	.030		.024			.036	.012	
13	13			.006		.006		.006	.012	.012		.006	.006	.779	.012		.012	.006	.006		.006
5	14	.012	.006	.006		.048		.006			.006	.012	.024		.839	.018	.006			.137	
8	15		.006			.012	.006	.018	.012	.006	.006	.006	.095	.048	.012	.833	.006		.012	.012	.012
6	16		.012	.006		.006	.024		.012	.012	.006	.012					.881	.012	.006		.006
19	17	.018	.018	.018	.006	.006	.006	.167	.006	.018			.006	.018	.006			.559		.024	.012
2	18		.006	.006	.006	.024	.006	.006	.018	.012		.006	.060		.012	.012	.006		.898	.030	
20	19	.012	.006	.012	.006	.006			.018	.006	.006		.012	.006	.060	.006	.006			.690	
3	20	.006	.006	.060	.024		.036	.024	.006	.125			.006	.012			.006	.030			.869

Operator Trends in Marking a Moving Scale Indicator

HAROLD R. WILLIS, Senior Systems Engineer, Aerospace Psychology Group, McDonnell Aircraft Corporation, St. Louis, Missouri.

Although the moving scale, fixed pointer, type of indicator is undesirable for normal use, occasions sometimes arise when it is appropriate, or its use is unavoidable. In one situation of this nature a dial required markings that would designate a quadrant on one side of 0 as North, and the other side as South. Another dial required markings on one side of 0 as Plus, and the other side as Minus. In the absence of specific information covering this problem, an experiment was designed which would supply the necessary information. A naive population of 40 randomly selected, male, senior high school students was used. Results showed that the right side of the moving scale should be marked N or + and the left side S or —. Peripheral information disclosed correlation between score obtained and scholastic performance.

INTRODUCTION

A considerable amount of research has been conducted to determine the best design for mechanical indicators. Their size, shape, scale dimensions and construction, pointers, direction of rotation, and orientation have been fairly well established[1, 2, 3, 4, 6, 8]. Among the problems that have been probed is that of the moving scale and the fixed pointer for which certain facts have been established[1, 6]. It has been found that some ambiguities exist with moving circular scales and associated control movement. When a moving circular scale with a fixed pointer is used, these recommended practices will invariably be violated:

1. Scale numbers should increase in a clockwise direction around the dial. Values on moving circular scales, therefore, increase with counter-clockwise rotation of the dial face.
2. The direction of movement of the associated control should be compatible with the direction of movement of the dial. (Clockwise movement of the control should result in clockwise movement of the dial.)
3. Clockwise movement of a control should result in an increase in function.

If Rule 1 is violated the operator will have a tendency to make errors in the final setting of the dial. A breach of Rule 2 finds the operator making errors in the initial direction in which the dial should be turned. Ignoring Rule 3 abrogates the standardized control movement vs system relationship.

Inasmuch as design criteria and space limitations occasionally make it necessary to use circular moving scales with a fixed pointer, recommendations have been made which are intended to make the best of an undesirable situation.

1. The numbers should progress in magnitude in a clockwise direction around the face of the dial such that when the dial is rotated counterclockwise, the reading indicated by the pointer will increase.
2. If a crank or knob is used to control the movement of the dial, the scale would move counterclockwise with a counterclockwise movement of the control, except in instances where the direction of movement of the control is directly related to the direction of movement of a vehicle. In this case the following should apply:
   a. The scale should rotate counterclockwise with a clockwise movement of a knob, wheel, or crank.
   b. The scale should rotate counterclockwise with movement of a lever forward, upward, or to the right.
   c. The scale should rotate counterclockwise with the movement forward, upward, or to the right of a vehicle or component.

3. The pointer should be centered at the 12 o'clock position for right-left directional indication and at the 9 o'clock position for up-down directional indication. Either position may be used to display strictly quantitative information[1].

Although such general rules as these have provided a criterion for direction of rotation and for progress of numbers around the dial, there are other problems in the use of moving scale, fixed pointer controls that require further examination and establishment of criteria.

Recently a missile was designed which required ground set dials that must indicate Plus and Minus on one dial and North and South on another dial. A moving scale, fixed pointer situation was mandatory because of the limited space in which the dials had to be located. The question arose as to which side of the dial should be marked Plus and which side Minus on the one, and which side should be marked North and which South on the other. An example of the design of the dials that are required is shown in Fig. 1.

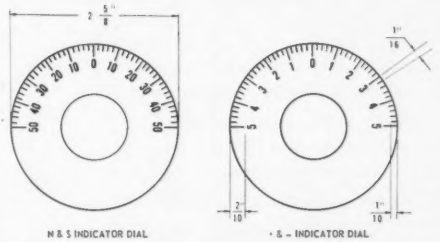

Fig. 1. Types of dial markings required in the missile.

A search of available literature disclosed only guidelines relative to how these dials should be marked[1, 5, 7]. As shown in Fig. 1, it is necessary that the dials be marked and numbered in both directions thereby eliminating the determination of which direction the scale should run. The major problem was to determine which side of the dials to establish as N & S and + and —. The following were points for consideration.

1. There is a general tendency to refer to North first and Plus first when speaking of them. Would it be desirable to follow this trend in moving from N to S and + to — as a left to right movement?
2. When we look at a 360° circle with North at the top, we move to the right around the scale in approaching S. Should we therefore follow this concept and place N on the left side of the scale and S on the right side?
3. If we do this, the movement of the scale under the pointer would show S for a counterclockwise movement of the scale, and N for a clockwise movement; or — and + respectively.

One of the major aspects of the problem of control and scale design is that of determining whether or not manipulation and reading by the operator will be "natural." The terms "natural" and "unnatural" are confusing in the context because of the dual reference to ease of discrimination and to unlearned modes of interpretation. This confusion can be dispelled by substituting the concept of a commonly observed behavior pattern or population stereotype. The latter concept implies only the notion of frequency of occurrence with no implication that the response is learned or innate[6].

According to Fitts[6] some writers have stated that even if an individual over-learns a response that is contrary to an earlier habit he will frequently revert to the population stereotype under stress of an emergency. There is no experimental evidence to show that this hypothesis holds for the interpretation of instruments or the operation of controls although anecdotal reports appear to support it[6]. Inasmuch as there is also no experimental evidence to the contrary the advisability of following the trend laid down by such anecdotal reports seems reasonable. Other studies have pointed out that adherence to the population stereotype in operation of controls achieves greater speed and accuracy and that training time can be minimized by taking advantage of stereotype response patterns, hence

providing an economy of learning[2]. The fact remains that population stereotypes for one group of people may not be the same for another. For example, it is "natural" for us to automatically throw an electric light switch "up" for "on" and "down" for "off" in the United States, but in Britain the opposite is the case.

METHOD

Determination of Apparent Population Stereotype

A series of tests was made to determine the apparent population stereotype that could be anticipated in an experiment designed to establish the marking of a rotating scale N & S and + & −. The dial chosen for the experiment was one that was similar to the North-South indicator being used in the missile. The dial is $2\frac{5}{8}$ inches in diameter and the control knob is 1 inch in diameter; both according to established standards[4]. The scale marking was designed according to tested and recommended procedures[1]. Twenty of these were produced.

In preliminary tests one dial was attached to a small panel and subjects were given several trials at setting the dial to various positions read off by the experimenter, e.g., North 25, South 17, Plus 15, Minus 32, and so on, to determine which direction the dial would be turned in the greater number of cases as the subject made his own selection. It was found that use of only one dial was unsatisfactory for the following reasons:

1. The operator had too much time to think about what he was doing to show what his "natural" trend might be in setting the dial.
2. Too many operations of the single dial were necessary to obtain the required information.
3. It was difficult to create a mild "stress" situation that would cause the operator to follow his natural tendencies rather than "think them out."

After considerable cut-and-try it was found that a panel with six dials would provide the desired material for the experiment. The dials were arranged as shown in Fig. 2.

A record sheet was designed, but it was modified three times before one was found that would best suit the purposes of the experiment. This is shown in Fig. 3.

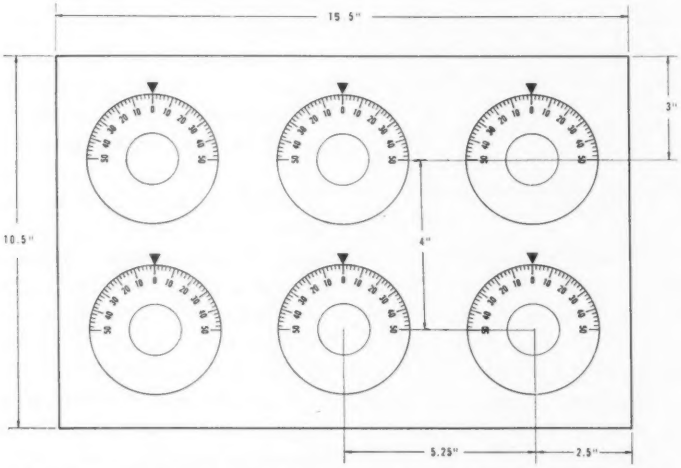

Fig. 2. Panel arrangement for moving scale indicator experiment.

MOVING SCALE RECORD

NO. _____

DATE _____

UNMARKED INDICATOR			MARKED INDICATOR		
REQ. SETTING	DIRECTION ERROR	SCALE ERROR	REQ. SETTING	DIRECTION ERROR	SCALE ERROR
N 16			N 16		
S 19			S 19		
S 32			S 32		
+ 23			+ 23		
N 35			N 35		
S 21			S 21		
– 38			– 38		
N 17			N 17		
– 13			– 13		
+ 43			+ 43		
S 18			S 18		
N 24			N 24		
– 41			– 41		
– 7			– 7		
+ 33			+ 33		
N 9			N 9		
+ 15			+ 15		
– 26			– 26		
S 8			S 8		
+ 41			+ 41		
N 32			N 32		
– 50			– 50		
S 17			S 17		
+ 19			+ 19		
+ 36			+ 36		
– 8			– 8		
N 47			N 47		
– 21			– 21		
S 33			S 33		
+ 14			+ 14		

NAME: AGE SEX

DEXTERITY PLANS: COLL. WORK SVC.

EXPERIENCE:

REMARKS:

EVALUATION:

Fig. 3. Moving scale record sheet used to record data for the experiment.

Preliminary trials with the six dials disclosed that if six settings were read in 10 seconds the subject would have to hurry enough to create a stress situation, but he would still be able to set the dials with a fair degree of accuracy.

It was also found that the subject took the most time on the last dial and got it correct most frequently because the degree of stress was more or less terminated at that point. In order to carry the stress situation through for all of the dials, it was decided that the subject would be told that he was being timed and he must tap a bell as soon as he finished the last dial so that the experimenter could tell when he was through. Additional trials disclosed that the desired stress through all dials was achieved.

Throughout the preliminary trials revision in the instructions to the subject was made until the following set of instructions was developed and found satisfactory for conducting the experiment:

"The object of this experiment is to determine how rapidly and accurately you can set these dials. Speed and accuracy are of equal importance. You will note that the scale runs from 0 to 50 to the left and 0 to 50 to the right around the dial. I will read a series of settings to you, such as, South 25, North 15, Minus 20, Plus 32 and so on. You will begin with the first dial and moving to the right; set a dial for each reading I give you. You will observe that the dials are not marked, North, South, Plus or Minus. When you make your settings you are to decide yourself which side is South, North, Minus or Plus. Once you have done this, try to be consistent.

"The six settings will be read in about 10 seconds, therefore, you must make each setting quickly and move on to the next one. Inasmuch as we are interested in 'speed,' I will try to pace you such that I am always one setting ahead of you, therefore, you may complete the setting in more or less than 10 seconds. I will time you. When you are finished with the last dial, tap the bell as quickly as you can so that I will know you are done. Are there any questions?"

If there were any questions, they were answered, then the subject was told; "Now I will read off six dial settings in 10 seconds so that you will know how fast you will have to move to complete the settings." The settings were then read off. This completed, the experimenter told the subject, "I will say 'ready,' then I will commence reading the settings." The idea of "speed" and "accuracy" was implanted in his mind in order to distract attention from the direction of rotation thereby aiding the "natural trend" to have more free expression.

A series of experiments was run on twenty subjects chosen at random from engineering, secretarial, and library personnel at McDonnell Aircraft Corporation to test the experimental procedure. No changes were found to be necessary with the exception of the third revision of the record sheet which was shown in Fig. 3. It was found, however, that the subjects were interested in knowing "how I made out," so a scoring system was set up as a motivating factor. The subject was scored as follows: one point off for each error of over one division in setting the dial, and two points off for each error in direction once the operator had established his selected direction for North, South, Plus and Minus. Some interesting facts were disclosed. On the average, a subject was found to make 13 errors during one run. It was also found that 77% of the N settings were made on the right side of the dial, and 73% of the S settings were on the left side of the dial. It was found that 82% of the + settings were made on the right side of the dial and 82% of the — settings on the left side of the dial; 75% of the subjects decided the right side of the dial should be N, the left side S, and 80% chose the right side for + and the left side for —.

The difference between percent of N and S and + and — selections made could be interpreted to mean that it was easier for the subjects to think in terms of + and — in relation to orientating them on the dials than it was of N and S.

These preliminary tests having been com-

pleted and the method of conducting the experiment established, arrangements were made to conduct the experiment for record upon a naive population, and at the same time one which would contain those individuals who would be the most likely to operate these dials in the armed forces. The population chosen was 40 male students who were seniors at the Maplewood-Richmond Heights High School.

Procedure for Conducting Experiment for Record

Preliminary

An evaluation of the information gained in the preliminary trials of this experiment showed that it would be desirable not only to determine the direction which the population was disposed to turn the dials for the N & S and + & − directions, but to determine errors made when they were also required to perform on dials that were clearly marked in order that some comparison could be made between performance on the two types. Two additional panels were therefore constructed similar to the one shown in Fig. 2. One panel was marked on the back with the identification letter "A," and the other panel with the letter "B." Panel "A" dials were marked with N & + on the right portion of each dial and S & − on the left portion. The "B" panel was marked with the N & + on the left portion of each dial, and S & − on the right portion. Preliminary trials were run on a test group with a subject first being tested on the panel with no markings, and then on an "A" panel; the second subject was tested on the panel with no markings and then on the "B" panel, and so on in an alternate series. With this accomplished satisfactorily, the final phase of the experiment was ready for execution.

Subjects

The subjects used in the final phase of the experiment, conducted at the Maplewood-Richmond Heights Senior High School at Maplewood, Missouri, were 40 male high school students in the senior class. Their ages were as follows:

"A" group—17 years old, 11; 18 years old, 9
"B" group—16 years old, 2; 17 years old, 12; 18 years old, 5; 19 years old, 1

The subjects were selected at random from physical education classes by the athletic coach and sent to the private room where the tests were conducted. The coach had been briefed previously on the experiments and what we hoped to accomplish. A later check with the high school principal disclosed that the random selection was satisfactory inasmuch as the subjects ranged from the highest to the lowest in scholastic performance as shown by their school records. None of the subjects was handicapped, 16 were athletically inclined, 6 were both athletes and scholars, 2 were scholars only, 16 were neither scholars or athletes.

Fourteen of the group expected to attend college after graduation. Of this group 7 were athletes, 4 were athletes and scholars, 3 were neither athletes or scholars. Two of the group of 40 were Negroes.

Apparatus

The material used consisted of the following: three panels containing 6 dials each mounted as shown in Fig. 2. One panel contained no N & S, or + & − markings; one panel designated as panel "A" had the dials marked with N & + on the right portion of the dial, and S & − on the left; one panel designated as panel "B" had the dials marked with S & − on the right portion of the dial, and N & + on the left; one stop watch; record sheets and pencils completed the apparatus.

Experimental Design

The final experimental design required that each subject perform on the unmarked panel, and then alternating, one subject would be tested on the "A" panel, and the next subject would be tested on the "B" panel, throughout the entire group. When the experiment was

completed all 40 subjects had been tested on the unmarked panel, 20 of them had also been tested on the "A" panel, and the other 20 had been tested on the "B" panel.

The directions for the subjects as previously outlined were given, then six settings were read off and, inasmuch as no bell was available to tap, the subject was required to rap the desk with his knuckles as soon as he had completed setting the last dial. Five groups of settings were read off for each subject making a total of 30 settings which each subject was required to make on each panel. Each group of six settings was timed, and the subject told how well he did. The experimenter watched the subject carefully out of the corner of his eye and kept one setting ahead of him as each dial was adjusted. At the end of each group of settings the experimenter checked the dials to determine whether or not the subject had made an error in direction of setting after he had selected N & S and + & −, and how much of a scale setting error had been made. This information, the time, direction of setting error, and scale error were recorded on the Moving Scale Record Sheet. When the subject had completed the series of five groups of six settings of unmarked dials, the experimenter selected an "A" or a "B" panel and instructed the subject that another series would be gone through. He explained that this time the dials would be marked, and that the subject should move the dials according to the way they were marked as the experimenter read off the settings. The same checking procedure was carried on for the marked dials.

The name of the student was recorded on the record sheet at the beginning of the experiment. After the experiment was completed he was asked to supply additional information as follows: age, whether or not he was right-, or left-, handed; his plans after graduation relative to attending college, working, or going into the armed forces; his work experience and hobbies. The experimenter then made a subjective evaluation of the performance of the subject stating whether he had appeared nervous during the experiment, whether or not he had adapted readily, his attitude toward the experiment, his competitive spirit, his interest in the experiment, his dexterity; the method by which he attacked the problem, i.e., emphasized speed at the expense of accuracy, or accuracy instead of speed; whether or not he appeared to be confused, and whether or not he had difficulty in understanding the instructions.

Students who asked what the experiment was about were told that to explain the experiment to them at this time would invalidate it, but that all who participated in it would be told about the purpose and the results at a meeting of the group at some future date. The experimenter returned to the school several weeks later and, during a home room period, briefed the 40 students on the purpose and the results of the experiment, and answered numerous questions.

RESULTS

Four out of the "A" group and four out of the "B" group chose to assign N & + to the left side of the dial and S & − to the right side, for an average of 1 in 5. None of these were, however, among students who turned in outstanding performances on direction and accuracy settings. The proficiency standing which they achieved among the 40 subjects was 7, 20, 21, 23, 25, 28, 31 and 35. The minority group of 8 represented 20 percent of the population participating in the experiment. Further evaluation showed that out of all errors made they were responsible for 20.8%. On the unmarked dials they made 15.5% of the direction errors, and 21.2% of the errors in setting; on the marked dials they made 24% of the direction errors, and 18% of the setting errors. In summary, they appear to be no better or no worse in performance than the majority group which chose N & + for the right side, and S & − for the left side of the dials.

Table I gives a breakdown of the number of times that a setting was made using the right or the left side for N or S or + or −. This

TABLE I

This Shows the Number of Times the Right Side of the Dial was Set on N or + and the Left Side on S or — Positions; "A" Group was Tested First on the Unmarked Panel, Then on the Marked "A" Panel. "B" Group was Tested First on the Unmarked Panel, Then on the Marked "B" Panel.*

	Right side north	*Left side south*	*Right side +*	*Left side —*	*Left side north*	*Right side south*	*Left side +*	*Right side —*	Misses
"A" Group Settings	90	88	105	105	43	48	52	51	18
"B" Group Settings	89	97	118	113	48	41	34	47	13
% of Settings by "A" Group	64	63	65	65	31	34	32.5	32	
% of Settings by "B" Group	63	69	74	70.6	34	29	21.3	29.4	
Total Settings	179	185	223	218	91	89	86	98	
Total % of Settings	64	66	69.8	68	32.5	31.8	26.9	30.6	

* "A" Panel was marked with N and + on the right side of the dial and S and — on the left side. "B" Panel was marked with N and + on the left side of the dial and S and — on the right side.

gives an appreciation of the distribution of settings that were made by both the normal selectees and the opposites regardless of the selection they had made. Eighteen misses by the "A" group and 13 by the "B" group in which the subject failed to set the dial at all accounts for the fact that the number of N & S settings does not add up to 560 and the + and — settings to 640 for the entire group on the unmarked panel.

Table II shows the errors that were made by each group on both the unmarked and the marked dials in direction and setting. There was a total of 1462 errors, or an average of about 36 per subject. Direction errors made on unmarked dials were 28.3% and setting errors 56.5%; direction errors on marked dials were 7.36% and setting errors were 48.5%. Total directional errors were 17.3%, and setting errors, 51.8%.

Table III shows a comparison between the performance of the engineering group used in the preliminary trials and the student group used in the experiment.

DISCUSSION

The results clearly indicated that the normal tendency will be for individuals to select the right side of the dial for North or Plus, and the left side of the dial for South or Minus by a ratio of 4 to 1 in instances where it is necessary to use moving scale, fixed pointer dials similar to those used in this experiment.

It may be assumed that those who would make the opposite choice could successfully carry on with this mode of dial marking according to evidence forthcoming from this experiment which showed that the minority group averaged approximately the same in direction and setting errors as did the majority group.

It may also be concluded from the data gathered that the selection of the right and left side of the dial as Plus and Minus, respectively, is easier and more natural to select and manipulate than a North and South selection.

The simple scoring system designed to motivate the subjects not only to do their best in the experiment, but to add to the stress has revealed some interesting peripheral information not relevant to the major problem of determining how the dials should be marked. The scores in the majority of instances compared favorably with the class records of the subjects. But in some instances the score of the subject was high but the class record was average or only slightly above average. Consultation with some of the school authorities revealed that in each case the

TABLE II

Errors Made on Both the Unmarked and the Marked Dials

	Unmarked dials						*Marked dials*					
	Direction errors			*Setting errors*			*Direction errors*			*Setting errors*		
	N	*S*	+	*N*	*S*	+	−	*N*	*S*	+	−	*N*	*S*	+	−
Number by "A" Group	22	27	25	20	87	93	94	89	6	6	4	1	78	78	88	84
Number by "B" Group	28	20	21	6	74	75	87	67	7	5	8	8	60	63	76	54
% by "A" Group	15.7	19.3	15.6	12.5	62.1	66.5	58.8	55.6	4.3	4.3	2.5	0.63	65.7	65.7	55	52.5
% by "B" Group	20	14.3	13.1	3.8	52.8	53.5	54.4	41.8	5.0	3.6	5.0	5.0	42.8	45	47.5	33.8
Total Errors	50	47	46	26	161	168	181	156	13	11	12	9	138	141	164	138
% Total Errors	17.9	16.8	14.4	8.1	57.5	60	56.5	48.6	4.6	3.9	3.7	2.8	49.2	50.2	51.2	43

Total direction errors	*Unmarked*	169	*Marked* 45
% Direction errors		28.3	7.5
Total setting errors		666	581
% Setting errors		56.6	48.5
Total direction errors both marked and unmarked		214	
% Direction errors both marked and unmarked		17.8	
Total setting errors both marked and unmarked		1247	
% Setting errors both marked and unmarked		51.8	

subject was believed to be capable of much more than he had demonstrated in class. In most instances, however, the individuals making high scores were also high in class standing. In no instance did students of low class standing score high on this test. Of the 12 subjects making a high score ranging from A− to A+, 3 were athletes and scholars, 2 were scholars, 4 were athletes, and 3 were neither scholars nor athletes. Of the 9 who scored low ranging from C− to D, 4 were athletes and 5 were neither athletes nor scholars. In one instance a boy known to be an outstanding student and an athlete made a C score. In the interview with him following the subjects performance, he stated that he did not like to work under pressure. A brief discussion with each individual after they had completed their two experimental runs revealed much about their impressions and attitudes. This information was recorded on the Moving Scale Record. It was interesting to note that 28 of the group either intended or desired to attend college.

It is believed that tests of performance such as this, properly designed, would come closer to actually evaluating the total capabilities of an individual than separate performance records, separate intelligence tests, separate achievement tests integrated for each individual. This was in the back of the mind of the originator when he designed this set of tests. It is felt that the tests as they were designed dealt with spatial relations in the handling of the dials and in the selection of the N & S or + & − side of the dials. Memory was also involved in remembering settings read off by the experimenter always far enough ahead of the individual's performance to require that he remember — and perform — under stress. Remembering his own selection and following it through, and his ability to remain calm and perform efficiently, accurately, and effectively under pressure demonstrated a composite of intelligence and performance that, in a crude way, gave a clue to the probability of the subjects' over-all capacity for intelligent performance under something more than classroom or laboratory conditions. This was considered

TABLE III

Comparison of Response and Selection by Engineering and Student Groups. Each Group Consisted of 20 Subjects.

	Engineering group	*Student group "A"*	*Student group "B"*	*Student total*
"N" Settings on Right Side of Dial	77%	64%	63%	63.5%
"S" Settings on Left Side of Dial	75%	63%	69%	66%
+ Settings on Right Side of Dial	82%	65%	74%	69.5%
− Settings on Left Side of Dial	82%	65%	70.6%	67.8%
Number of Subjects Selecting Right Side for "N" and Left for "S"	75%	80%	80%	80%
Number of Subjects Selecting Right Side for "+" and Left for "−"	80%	80%	80%	80%

valuable information inasmuch as similar dials would be used by armed forces personnel. It is suggested that a more vital figure than I.Q., which has been held in such high esteem for so many years, would be the P.Q., or Performance Quotient, that included not only the capacity for learning, but gives an indication of the individual's ability to apply his intelligence to life situations.

REFERENCES

1. BAKER, C. A., and GRETHER, W. F., *Visual Presentation of Information.* USAF WADC Tech Rep 54–160, August 1954.

2. CHAPANIS, A., GARNER, W. R., and MORGAN, C. T., *Applied Experimental Psychology,* 1949. Wiley: New York.

3. ELY, J. H., THOMSON, R. M., and ORLANSKY, J., *Layout of Work Places.* USAF WADC Tech Rep 56–171, September 1956.

4. ELY, J. H., THOMSON, R. M., and ORLANSKY, J., *Design of Controls.* USAF WADC Tech Rep 56–172, November 1956.

5. ELY, J. H., BOWEN, H. M., and ORLANSKY, J., *Man Machine Dynamics.* USAF WADC Tech Rep 57–582, November 1957.

6. FITTS, PAUL M., *Psychological Aspects of Equipment Design.* USAF WADC Tech Rep No. 5829, August 1949.

7. MCCORMICK, E. J., *Human Engineering,* 1957. McGraw-Hill: New York.

8. WOODSON, W. E., *Human Engineering Guide for Equipment Designers.* University of California, Berkeley, 1956.

Book Review

PSYCHOLOGY AND HUMAN PERFORMANCE: An Introduction to Psychology by Robert M. Gagné and Edwin A. Fleishman

Henry Holt and Company, Inc. New York 1959 xiii+494 pp. Price $7.25.

It is sometimes hard to convince an engineer or a physiologist that his background may have serious holes in it with respect to the fundamentals of psychology. To help solve this problem, an elementary text on psychology that would be suitable for orienting non-psychologically trained human factors workers to the basic concepts of psychology has been much needed. It has been the experience of this reviewer that the typical elementary texts in psychology which have been available over the past years, have not had sufficient relationship to practical applications with which they are familiar to convince non-psychologists that they would profit by time spent in the study of elementary psychology. *Psychology and Human Performance* presents an approach which emphasizes the tremendously expanded knowledge that has occurred during and since World War II in the field of motor skills. It thus may be a text in introductory psychology that will get widespread support from human factors personnel.

Because of the emphasis on human performance the book warrants the interest of both psychologists and non-psychologists as an introductory text which covers the ground rules that govern applied decisions. On page vi of their preface the authors state: "Although fundamental principles are stressed throughout, there is also a practical emphasis to this book. At least implicitly, it offers the point of view that psychological knowledge can and should be used in the planning and design of productive work enterprises that require the participation of human skills. In the final four chapters, this idea is given an expanded treatment in connection with human jobs, personnel selection, proficiency measurement, training programs, and the design of equipment and man-machine systems." Both the authors, now in academic institutions, have had extensive experience in applied fields; Gagné as technical director of a government laboratory for several years, and Fleishman as a scientist with the Air Force Personnel and Training Research Center, Lackland Air Force Base. The organization of the book and the selection of both verbal and pictorial illustrations reflect heavily the contact which the authors have had with applied research. The extensive experience which Fleishman has had in the factor analysis of motor skill components is drawn on in the context of many other related studies to make Chapter 5 on human abilities a unique contribution to an introductory text book. Chapter 8 on motor skills will also be found to be a considerable departure from the approaches used in other recent elementary texts. In this chapter as well as in several others, particularly the last four, selected studies that will probably long remain as classics for the psychological aspects of human factors engineering are reviewed. No professionally qualified human factors worker can fail to feel "at home" with the many references in the book which are drawn from those of his own "bread and butter" daily contacts. Because of its references to the literature of applied research *Psychology and Human Performance* appears suitable for use as one of the standard texts for orienting neophytes at the lower division undergraduate college level as well as professionally advanced non-psychologist newcomers to the field of human factors and its realizable promises. If the text becomes popular for introductory courses in psychology, it should not fail to help increase the number of people who wish to seriously pursue human factors specialties professionally.

From the standpoint of presentation *Psychol-*

ogy and Human Performance is lucid and well organized even though it tends to be somewhat superficial in its treatment of many topics such as the physiological basis of the senses and effectors. Its authors have also tended to select pictorial illustrations (e.g. Figure 14.2, Page 435) uncritically. At the level of training for which the book is intended neither fault will probably prove to be a special handicap in the face of the strong and competent description of overall human performance as related to experimental research in the field of perceptual and motor skills.

The available facts are adhered to closely and there is a minimum of speculative interpretation. Each chapter is preceded by a statement of its intent and direction and followed by a summary of its content. The main chapter headings are as follows:

1. Introduction: The Science of Behavior
2. Basic Components of Behavior
3. Functions of the Behavior System
4. Motivation
5. Human Abilities
6. Learning and Retention
7. Discrimination and Identification
8. Motor Skills
9. Concepts and Thinking
10. Social Behavior
11. Jobs and Personnel Selection, I
12. Jobs and Personnel Selection, II
13. Training
14. Engineering Psychology

The book gives the overall impression of a pertinent and high quality contribution to the library of available materials for human factors personnel. Its content is an excellent demonstration of how matters that once were ideas at the frontiers of our knowledge may later become part of the common content of introductory courses at the freshman-sophomore college level.

John Lyman
UCLA

The Authors

THEODORE W. FORBES (*Human Factors in Highway Design, Operation and Safety Problems*) is the assistant director in charge of research and professor of psychology at Michigan State University's Highway Traffic Safety Center. He directs and coordinates work in traffic engineering, economics, sociology and psychology as related to traffic operations.

For the past 25 years Dr. Forbes has been a teacher, lecturer and researcher specializing in highway traffic psychological and engineering problems, and has done considerable work under the Office of Scientific Research.

He was born in Honolulu, Hawaii. He received his A.B. (mathematics, physics and sociology) in 1924 and his M.A. (physiological psychology) in 1927 from Oberlin College. His Ph.D. (applied experimental psychology) was received from Ohio State in 1931. He was certified in traffic engineering at the Harvard Traffic Bureau in 1938.

Dr. Forbes has held positions with Ohio State University, the New York State Psychiatric Institute, Barnard College of Columbia University, the Harvard and the Yale Traffic Bureaus, Stevens Institute of Technology, University of Pennsylvania, Harvard Psycho-Acoustic Laboratory, University of Hawaii and the University of California at Los Angeles. Prior to his post as program director in charge of highway research at the American Institute for Research, Dr. Forbes was technical director of the Committee for Highway Safety Research, National Academy of Sciences, National Research Council.

Pertinent research activities and publications include highway sign legibility, driver tests, accident analysis, auditory signals for instrument flying, auditory research and human engineering and psychological applications for traffic engineering and highway safety.

He is an associate of the Highway Research Board and chairman of the committee on road user characteristics, a member of the President's Traffic Safety Committee; and is a fellow in the American Academy for the Advancement of Science and the American Psychological Association. He is a member of a number of other professional organizations, including the Human Factors Society.

ELVIN E. HERMAN (*Prospectus for Human Factors in Electronics*) is Manager of Advanced Development, Hughes Aircraft Company, Culver City, California. He is responsible for the technical and administrative direction of radar analysis, human factors, and applied research and development in electronic sensing and data processing systems.

Mr. Herman received his B.S. in electrical engineering from the University of Iowa in 1942. After graduation he spent 6 years at the Naval Research Laboratory primarily in electronics counter-countermeasures and naval search radar research and development. In 1951 he transferred to the National Bureau of Standards and was responsible for the design, development and evaluation of missile guidance equipment. In 1953 he joined the Airborne System Laboratories of the Hughes Aircraft Company. He specialized in planning radar systems, interceptor-testing of fire control equipment and applied research and development of display storage tubes and radar signal processing.

He is a senior member of IRE, a member of Eta Kappa Nu and has received the Navy Meritorious Civilian Award. He has also been issued a number of patents in electronics.

RAYMOND E. BERNBERG (*A Comparison of Three Flight Attitude Displays*) is presently Supervisor, Engineering Psychology, Human Factors Group, at North American Aviation, Inc., Los Angeles Division. This group is responsible for crew stations design in aircraft and manned weapon systems. Two major weapon systems with which they are presently concerned

are the B-70, Mach 3 SAC Bomber and the F-108, Mach 3 Long Range Interceptor for ADC.

Dr. Bernberg received his B.A., in 1948, his M.A., in 1949 and his Ph.D., in 1950 in psychology from the University of California, Los Angeles. Prior to joining North American Aviation, Inc., he was an associate professor at Los Angeles State College and a consultant to many Southern California missile, airframe and electronics firms.

JOHN J. DREHER and WILLIAM E. EVANS (*Speech Interference Level and Aircraft Comfort*).

JOHN J. DREHER is a research specialist in the Physics Research Department at Lockheed Aircraft Corporation, Burbank, California. He is also a lecturer in the graduate program at the University of Southern California.

A military specialist in Chinese during World War II, he completed his Ph.D. in experimental phonetics at the University of Michigan in 1951, while serving on the faculty at that institution. He served two summers as a research associate at the U.S. Naval School of Aviation Medicine, Pensacola, Florida. In 1951 he joined the Ohio State University Research Foundation as principal investigator at the Psycholinguistics Laboratory, where he worked on air traffic control phraseologies for International Language of the Air. Dr. Dreher is a member of several societies, and is the author of more than 50 technical papers.

WILLIAM E. EVANS is a research engineer in the Physics Research Department at Lockheed Aircraft Corporation, Burbank, California. Prior to joining Lockheed, he was associated with the Acoustics Section of Douglas Aircraft Company, Santa Monica, after being released from active duty as a communications officer in 1956. While at Douglas, he worked on psychoacoustic and communications problems in both aircraft and missile systems.

Mr. Evans received his M.A. in psychology and speech science at Ohio State University in 1954 while serving as a Research Assistant for the Ohio State University Research Foundation. Concurrently he worked on visual and oral communications problems for the U.S. Navy Voice Communications Laboratory at Pensacola, Florida. He is a member of the Human Factors Society and the Acoustical Society of America and is currently engaged in a Ph.D. program at the University of Southern California.

HUGH M. BOWEN, JOHN L. ANDREASSI, SHAFFER TRUAX, and JESSE ORLANSKY (*Optimum Symbols for Radar Display*).

HUGH M. BOWEN has been concerned with radar display and interpretation problems since joining the Applied Psychology Research Unit, Cambridge, England, in 1951. He acted in a consultant capacity on radar and associated problems to the R.A.F. as a member of the Flying Personnel Research Committee. After taking his Ph.D. degree in 1956 from Cambridge University, he joined Dunlap and Associates, Inc., and is currently a research associate undertaking experimental and analytical work in the field of man-machine systems. In addition to activity in the visual and perceptual fields, he is undertaking projects in the analysis and training of motor control skills. He is a member of the Society of Engineering Psychologists, a Division of the APA. Recent publications include joint authorship of Chapter VII, "Man-Machine Dynamics" Joint Services Engineering Guide.

JOHN L. ANDREASSI graduated from City College of New York in 1956 with a Bachelor's degree in industrial psychology. He obtained his Master's degree in experimental psychology from Fordham University in 1959. Since joining Dunlap and Associates, Inc., in 1958, Mr. Andreassi has been engaged in a number of human engineering projects concerning visual displays, training methods and equipment design. He is currently concerned with human factors studies and investigations of Signal Corps equipment and systems. He is a member of Psi Chi and the American Psychological Association.

SHAFFER TRUAX has undertaken a number of projects in the analysis of display devices. After receiving his M.S. from City College of New York in 1953, he taught and was head of the Psychology Department at Bloomfield College. He left in 1956 to go to Rutgers University, getting his Ph.D. in 1959. He joined Dunlap and Associates Inc., in 1959 as an associate psychologist. He has undertaken studies of the effectiveness of warning signals while tracking and is currently engaged in examining perceptual recognition of infra-red images.

JESSE ORLANSKY has been a Vice President of Dunlap and Associates, Inc., since its start in 1948. He received his Ph.D. in psychology from Columbia in 1940, and until 1943 was an instructor in psychology at CCNY. During the war he was in the Navy where he worked in high altitude chambers and conducted research on aircraft lighting, camouflage, and night vision. At the Psychological Corporation (1946–1948) and in his present affiliation, he has worked on aircraft instumentation, maps, data processing centers, and ground support equipment. He is a Diplomate in Industrial Psychology, American Board of Examiners in Professional Psychology, and has served as consultant to Working Group VI, Armed Forces NRC Committee on Vision; the Psychology and Social Science Panel of the USAF Scientific Advisory Board; and the Psychology and Social Science Division, Department of Defense.

HAROLD R. WILLIS (*Operator Trends in Marking a Moving-Scale Indicator*) is presently a senior systems engineer in the Aerospace Psychology Group at McDonnell Aircraft. Born in Trenton, Missouri in 1910, he enlisted in the U.S. Navy at seventeen and retired as Commander in 1957. As Training Department Head at Fleet Airborne Electronics Training Unit, San Diego, he pioneered, from 1945 to 1949, in the development of Radar Target Intelligence for the U.S. Navy. He served as Editor of the *Digest of U.S. Naval Aviation Electronics* in Washington, D.C., from 1949 to 1952. He served with the Radar Target Intelligence Offices, Staff, Commander in Chief, Pacific Fleet, from 1952 to 1955. From 1955 to 1957 he served as Executive Officer U. S. Naval Photographic Interpretation Center, in Washington, D.C., as Executive Officer.

Commander Willis received his A.B. in education from George Washington University, Washington, D.C. He obtained his M.A. degree in education and his M.S. in psychology from Jackson College, Honolulu, T.H. He is now working toward the Ph.D. in experimental psychology at Washington University, St. Louis, Missouri. He is a member of Phi Delta Kappa, educational professional fraternity, The Human Factors Society, The Southern Association of Philosophy and Psychology and the Missouri Psychological Association.

Editorial

VOLUME TWO

At the beginning of a new volume, we wonder what we shall be able to do with it. We have some ideas, based on past submissions. We have some fears, and we have some hopes.

Our past submissions reflect a random sampling of information useful to the solution of human factors problems.

We have some fears that there is, unwittingly, a divisive trend that will lead us away from becoming an integrated, scientific discipline. This trend has two aspects, one of which is to seek a haven in the disciplines of nurture: electronics, mechanical engineering, psychology or physiology. The other is to regard human factors as an adjunct to automotive, submarine, aircraft or space system designs.

We have some hopes that gradually the miscellaneous trinkets of thought we have used in the past will give way to a unified and sound intellectual currency and that we shall at least bequeath a sound medium of exchange to the next generation of workers.

This journal must depend primarily on the members of the Human Factors Society for its literary lifeblood. Only as we work diligently to discover principles shall we be able to report principles. Only as we make extensive reviews can we publish such valuable contributions. Only as we set more rigorous standards for ourselves can we challenge our colleagues to higher levels of professional reporting. Only by spending a little of our precious time can we offer clarifying comments through letters to the editor.

We have over 500 members. If 10% are not contemplating a paper in the next twelve months, we will not have sufficient articles for the year. If we wonder why some valuable work of a colleague has not been published, the answer is that either he has not written it up, or he has not written it up well enough, or he has not submitted it for publication.

Each of us in our daily work is somewhat analogous to a primitive who brings a commemorative stone to a site. From these stones, we hope to erect a temple of thought. At the risk of being covered by a mountain of mail, your editor volunteers to receive bids for the construction.

Stanley Lippert

EXCELLENT OPPORTUNITIES FOR

PSYCHOLOGISTS

You are invited to investigate unusual career opportunities with excellent futures in the science and engineering Laboratories of IBM. You will find this to be a unique opportunity to contribute original thinking on several important projects.

Our Laboratories are engaged in the research and development of advanced electronic digital computers for business, science and industry. We are also engaged in the design of control systems for exotic-fueled strategic bombers to operate at altitudes above 70,000 feet and at speeds of more than 2,000 miles per hour.

- To undertake design, research, experimental analysis, and consultation in the field of man-machine relationships in a newly organized psychology group. Unusual opportunities for display and systems design in the early development stage. Projects include work in psychophysics, communications, electronics, trouble-shooting techniques, human learning, perception, operations analysis, high-speed data processing, psychometrics, information theory, decision processes, visual and auditory displays, and equipment design. Program includes both applied and basic experimentation. An advanced degree is required, or considerable experience in appropriate fields.

- To conduct systems analysis on complex, advanced weapons systems. A strong background in mathematics and statistics is desirable. An M.A. or Ph.D. in Psychology is required with experience in human engineering and human factors research.

- To initiate human factors design on Navy programs. Experience either on naval systems or equipment is highly desirable. An M.A. or Ph.D. in experimental psychology is required.

- To perform equipment evaluation experimentation, laboratory simulation, and flight test analysis from the human factors standpoint. An M.A. or Ph.D. in Psychology is required plus applied experimental experience.

- PERSONNEL AND TRAINING PSYCHOLOGIST to conduct maintenance and handling task-equipment analysis and other personnel analyses on advanced electronic systems. Knowledge of Air Force skill classifications and technical training and related equipment is desirable. Must have knowledge of airborne weapons system research and development cycle. An M.A. or Ph.D. in Psychology is required with specialization in industrial and engineering applications of psychological knowledge and techniques.

• • •

Salary commensurate with experience. Advancement based on merit. Employee benefits are among the best in industry today.

Please write, including resume, to:

Mr. R. E. Rodgers, Dept. 700 J
International Business Machines Corp.
590 Madison Avenue
New York 22, N. Y.

IBM®

INTERNATIONAL BUSINESS
MACHINES CORPORATION